DNA methylation in cancer development
Lessons learned from tumors of the testis and the large bowel

Guro Elisabeth Lind

Group of Molecular Cancer Genetics
Department of Genetics
Institute for Cancer Research
The Norwegian Radium Hospital

Faculty Division
The Norwegian Radium Hospital
Faculty of Medicine
University of Oslo

Norwegian Cancer Society

Oslo 2005, Doctor philosophiae thesis
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>5</td>
</tr>
<tr>
<td>PREFACE – “DNA: THE SECRET OF LIFE”</td>
<td>6</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>7</td>
</tr>
<tr>
<td>LIST OF PAPERS</td>
<td>8</td>
</tr>
<tr>
<td>GENERAL INTRODUCTION</td>
<td>9</td>
</tr>
<tr>
<td>Cancer as a genetic and epigenetic disease</td>
<td>9</td>
</tr>
<tr>
<td>Epigenetics</td>
<td>11</td>
</tr>
<tr>
<td>DNA Methylation</td>
<td></td>
</tr>
<tr>
<td>DNA methylation as a gene silencing mechanism</td>
<td>12</td>
</tr>
<tr>
<td>Chromatin</td>
<td>13</td>
</tr>
<tr>
<td>Components of the main chromatin-modification and -remodelling complexes</td>
<td>14</td>
</tr>
<tr>
<td>Epigenetic regulation of gene expression – an interplay of multiple components</td>
<td>17</td>
</tr>
<tr>
<td>DNA methylation in normal development</td>
<td>19</td>
</tr>
<tr>
<td>Genomic imprinting</td>
<td>20</td>
</tr>
<tr>
<td>X-chromosome inactivation</td>
<td>22</td>
</tr>
<tr>
<td>Tissue-specific genes</td>
<td>23</td>
</tr>
<tr>
<td>Germ line specific genes</td>
<td>24</td>
</tr>
<tr>
<td>Intragenomic parasites and repetitive sequences</td>
<td>25</td>
</tr>
<tr>
<td>DNA methylation in cancer</td>
<td>26</td>
</tr>
<tr>
<td>Hypomethylation, genomic instability, gene activation, and loss of imprinting</td>
<td>26</td>
</tr>
<tr>
<td>Hypermethylation and gene silencing</td>
<td>29</td>
</tr>
<tr>
<td>DNA methylation, ageing and diet</td>
<td>30</td>
</tr>
<tr>
<td>DNA methylation and mutations</td>
<td>31</td>
</tr>
<tr>
<td>Testicular germ cell tumors – a curable rare malignancy arising in young men</td>
<td>32</td>
</tr>
<tr>
<td>Incidence, treatment, and outcome</td>
<td>32</td>
</tr>
<tr>
<td>Histopathology</td>
<td>34</td>
</tr>
<tr>
<td>Genetic changes</td>
<td>34</td>
</tr>
<tr>
<td>Epigenetic changes</td>
<td>35</td>
</tr>
<tr>
<td>Colorectal cancer – a common malignancy with poor outcome</td>
<td>38</td>
</tr>
<tr>
<td>Incidence, treatment, and outcome</td>
<td>38</td>
</tr>
<tr>
<td>Histopathology</td>
<td>39</td>
</tr>
</tbody>
</table>
The adenoma-carcinoma sequence ................................................................. 39
Genomic instability ......................................................................................... 40
Epigenetic changes ......................................................................................... 41
CIMP or CIMPlly not? ...................................................................................... 45
AIMS OF THE PROJECT .................................................................................. 46
RESULTS IN BRIEF .......................................................................................... 47
DISCUSSION ..................................................................................................... 53
Methodological considerations ................................................................. 53
Modification of DNA - Sodium bisulphite treatment ........................................ 53
Methylation specific-PCR .............................................................................. 54
Interpretation of MSP results ....................................................................... 55
Bisulphite sequencing .................................................................................... 57
Validation of the effect of DNA hypermethylation ........................................ 59
Genome wide gene expression analyses in cell lines cultured with and without
demethylating agents .................................................................................... 60
DNA methylation in cancer development .................................................... 62
DNA methylation changes in testicular germ cell tumorigenesis .................. 62
DNA methylation changes in colorectal tumorigenesis ................................. 64
Disparities and similarities in epigenetic changes in TGCTs and colorectal carcinomas 67
Clinical applications of epigenetic changes in cancer .................................. 68
Gene promoter hypermethylation as a biomarker for early detection, intervention, and
prognosis ........................................................................................................ 69
Epigenetic silencing mediated by CpG island hypermethylation as a potential therapeutic
target .............................................................................................................. 70
CONCLUSIONS ............................................................................................... 73
FUTURE PERSPECTIVES ................................................................................. 76
REFERENCES .................................................................................................. 80
ORIGINAL ARTICLES ....................................................................................... 105
APPENDICES ................................................................................................. 111
Appendix I. General abbreviations ............................................................. 112
Appendix II. Gene symbols .......................................................................... 113
Appendix III. Glossary .................................................................................. 118
Acknowledgements

The work presented in this thesis has been carried out at the Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital from 2001-2005 and has been financially supported by the Cancer Society.

First of all, I will thank my supervisor Professor Ragnhild A. Lothe for letting me join her excellent research group. Her knowledge is exceeding far beyond her own field of research and working with her is truly inspiring. I especially appreciate the challenges she has given me, her constructive criticism, and our frequent scientific discussions.

I want to thank Dr. Manel Esteller for inviting me to his epigenetic facilities at CNIO in Madrid and his research group for taking good care of me, both scientifically and socially during my two-months stay. I owe special thanks to Dr. Mario Fraga for making epigenetic obstacles smaller, and to Lidia and Paula for letting me be the “third twin”. Further, I want to credit all collaborators and co-authors for optimal biologic material, clinical data, accurate pathology, and scientific input.

I also acknowledge the Head of Department, Professor Anne-Lise Børresen-Dale, for providing good research facilities and Professor Jahn Nesland, for being my official link to the Faculty of Medicine.

A special thanks goes to all present and former members of Ragnhilds’ research group, Rolf, Lin, Maja, Chieu, Kristine, Trude, Merete, Mette, Terje, Marianne, Helge, and Stine along with my friends at the department for creating a great scientific and social environment.

Last, but not least, I would like to thank my family and friends. You mean a lot to me.

Oslo, July 2005

Guro Elisabeth Lind
On February the 28th 1953, James Watson and Francis Crick deciphered the structure of deoxyribonucleic acid, DNA\(^1\), which was later published in the journal Nature [1]. The DNA X-ray diffraction pictures made by Rosalind Franklin were essential for this discovery, as they provided several of the vital helical parameters [2]. The four separate building blocks of the DNA molecule, the nucleotides adenine, cytosine, guanine, and thymine, had been isolated and characterized several years before this immensely important discovery. The modified cytosine base, 5-methylcytosine, was first recognized in 1948 [3] and was later identified as a central element in the field of epigenetics\(^2\). The DNA contains the genetic instruction specifying how to assemble protein molecules, which are the building blocks of each phenotype. Indeed, Crick described the DNA molecule as “the secret of life”, and today several fields of research address DNA directly or indirectly. The most recent breakthrough in the history of DNA research has been the sequencing of the human genome [4,5], which has heralded a new era for genetic as well as epigenetic research. The challenge now, is to understand the molecular mechanisms that allow specific genes and gene families to be selectively expressed in normal development and how aberrations in this process can lead to disease. In addition to well-described genetic mechanisms, imbalances in the epigenetic control of gene expression can profoundly alter this finely tuned machinery. Epigenetic changes are now recognized to have a lead role in cancer development [6]. Simultaneously, such changes have been hypothesized to be a master key to more effective ways of diagnosing, monitoring, and treating cancer [7]. On our way to molecular assisted medicine, we need to explore this in detail in order to get a better understanding of the role of epigenetics in cancer development, which is necessary to fully master these new tools.

\(^1\) See Appendix I for complete list of abbreviations.
\(^2\) See Appendix III for glossary.
SUMMARY

Through six reports, the current thesis investigates DNA methylation in cancer development, by exploring this phenomenon in two cancer diseases, testicular cancer and colorectal cancer. This study shows that the two cancer types display different methylation profiles and identifies novel gene targets inactivated by DNA hypermethylation during the tumorigenesis in each of the two organs as well as in their in vitro models.

In TGCTs, CpG island promoter hypermethylation is associated with tumor histology. Nonseminomas display methylation frequencies of target genes comparable to those of other cancers whereas seminomas in general are devoid of methylation. Among the various nonseminomatous histological subgroups, teratomas display more and embryonal carcinomas less methylation than do the remaining nonseminomatous subtypes. The methylation frequencies vary along the embryonic and extra-embryonic differentiation lineages of the testis tumor model, mimicking the epigenetic reprogramming during early embryogenesis. Epigenetic target genes novel in testicular tumorigenesis are identified, including MGMT, SCGB3A1 (HIN-1), and HOXA9.

DNA hypermethylation is found in small benign lesions of the colorectum and typically increases in frequency as the tumor progresses into malignancy. Apparently, right-sided hyperplastic polyps represent the exception, with a methylation level comparable to that of the carcinomas. The following genes, ADAMTS1, CRABP1, HOXA9, and NR3C1, are identified as novel epigenetically regulated target genes in colorectal tumorigenesis. Several associations were found among clinicopathological variables and gene methylation, which may aid in the future molecular assisted classification of colorectal cancer patients. Furthermore, colon cancer cell lines were found to be representative epigenetic models for the in vivo situation.
LIST OF PAPERS


III **GE Lind**, RI Skotheim, MF Fraga, VM Abeler, M Esteller, and RA Lothe. Epigenetically deregulated genes novel in testicular cancer development include homeobox genes and *SCGB3A1 (HIN-1)*. Submitted manuscript


V **Guro E. Lind**, Kristine Kleivi, Trude H. Aagesen, Gunn I. Meling, Torleiv O. Rognum, and Ragnhild A. Lothe. Identification of Novel Gene Targets in Colorectal Tumorigenesis: *ADAMTS1*, *CRABP1*, and *NR3C1*. Submitted manuscript

VI Terje Ahlquist, **Guro E. Lind**, Espen Thiis-Evensen, Gunn I. Meling, Torleiv O. Rognum, Morten Vatn, and Ragnhild A. Lothe. DNA methylation in benign and malignant large bowel tumors and their *in vitro* models. Manuscript
Cancer as a genetic and epigenetic disease

The development of cancer is a multistep process. Organisms are maintained by homeostasis, a finely tuned balance between cell proliferation and cell death. When the homeostasis is disturbed, either by an increased proliferation rate or a decrease in cell death, a neoplasm might occur, which can further progress into a tumor. Tumor development is most commonly described as natural selection followed by clonal expansion, resulting in monoclonal tumors originating from the progeny of a single cell [8]. However, cytogenetic studies indicating polyclonality have also been reported [9]. Aberrations that confer growth advantages to the cell will accumulate during the clonal selection process. These changes are consequences of 1) activation of proto-oncogenes, rendering the gene constitutively active or active under conditions in which the wild type gene is not, 2) inactivation of tumor suppressor genes, reducing or abolishing the activity of the gene product, 3) alteration of repair genes, which normally keep genetic alterations to a minimum [10].

Genomic analyses focusing on structural and numerical aberrations of chromosomes have long suggested that cancer is, in essence, a genetic disease [10]. The first cancer-specific genetic aberration described was the Philadelphia chromosome in patients with chronic myeloid leukemia. This was initially identified in 1960 by Nowell and Hungerford and was later demonstrated to be the result of a translocation between chromosomes 9 and 22 [11]. A synthetic drug, targeting the inappropriately activated gene product from this translocation, has recently been developed and is now administered to patients with the Philadelphia chromosome [12]. This successfully tailored treatment underline the potential for the development of
anticancer drugs based on the specific molecular abnormality present in a human cancer. Today, numerous mutations at the chromosome and DNA level have been described in hematological as well as solid tumors [10,13]. The Mitelman Database of Chromosome Aberrations in Cancer lists the chromosomal aberrations of more than 47,000\(^3\) tumors [14], and the IARC mutation database have recorded 21,587\(^4\) somatic mutations of the tumor suppressor gene \textit{TP53}\(^5\) [15].

During the last decades, several lines of evidence have proven the importance also of epigenetic modifications in tumorigenesis. Indeed, epigenetic changes are now recognized to be at least as common as genetic changes in cancer [6]. Moreover, epigenetic changes often precede and appear to be essential for several genetic events that drive tumor progression. Epigenetic inactivation of key genes in tumorigenesis, like \textit{p16\(^{INK4a}\)}, is seen in pre-malignant stages [16] and can allow cells to bypass the cell cycle restriction point, setting the stage for accumulation of more aberrations. A more direct link is seen with the epigenetic inactivation of the mismatch repair gene \textit{MLH1} leading to a microsatellite unstable phenotype of genome-wide insertions and deletions typically found in short nucleotide repeats [17]. Such repeats are also present in coding parts of the genome and mutations in cancer-critical genes, like \textit{TGFBR2}, may lead to a selective proliferation advantage for the cell [18-20].

---


\(^5\) See Appendix II for complete list of gene symbols and gene names approved by the Human Gene Nomenclature Committee; http://www.gene.ucl.ac.uk/nomenclature/
**Epigenetics**

Conrad Waddington introduced the term “epigenetics” in the 1940s to describe “the interactions of genes with their environment, which bring the phenotype into being” [21]. This early usage of the term has been effectively displaced during the last decades and today epigenetic inheritance is defined as cellular information, other than the DNA sequence itself, that is heritable during cell division [22]. Epigenetics affect the transcription in the cell, thereby controlling gene expression and abnormal epigenetic changes can have serious effects for the organism. We can very roughly divide epigenetics into three substantially overlapping categories: DNA methylation, genomic imprinting, and histone modification. Among these mechanisms, DNA methylation is the most studied, and is the main focus of this thesis.

**DNA Methylation**

DNA methylation is a covalent modification of nucleotides and the most frequently methylated nucleotide in the human genome is cytosine subsequently followed by a guanine in the DNA sequence, constituting a CpG dinucleotide. The cytosine is methylated in the C-5 position by a family of DNA (cytosine-5) methyltransferases (DNMTs) using the universal methyl donor S-adenosyl-L-methionine (SAM). 5-methylcytosines account for about 1% of total DNA bases in the human genome and affects 70-80% of the CpG sites in a human somatic cell [23]. Spontaneous deamination of 5-methylcytosine to thymine [24] has during the evolution led to a great under-representation of CpG dinucleotides in the human genome. When unmethylated cytosine deaminates to uracil it will be excised by the enzyme DNA-uracil glycosylase, and the original sequence is restored by DNA repair enzymes. However, the DNA repair machinery does not recognize the thymines resulting from 5-methylcytosine deaminations. Hence, the spontaneous deamination of 5-methylcytosine leads to a C to T transition mutation in the
genome [24] (see figure 1). A substantial fraction of the CpG dinucleotides left in the genome are located in CpG islands, which are GC-rich regions that possess high relative densities of CpG. They are mainly positioned at the 5' ends of many human genes and are usually unmethylated regardless of the expression status of the associated gene [25]. Recent reports using computational analyses suggests that there are at least 29,000 CpG islands in the human genome [5,26]. Several CpG sites are also found within repetitive or parasitic intragenomic elements. In contrast to the CpG islands, the CpG sites located here are largely methylated [27], as are the majority of the remaining CpG sites scattered in the human genome.

**Figure 1. Introduction of mutation by spontaneous deamination of 5-methylcytosine.** Deaminated cytosine resembles uracil and is excised by the enzyme DNA-uracil glycosylase and the original sequence is restored. Deaminated 5-methylcytosine resembles thymine and is left unrepai. This introduces a mutation in the genome.

**DNA methylation as a gene silencing mechanism**

The first connection between DNA methylation and gene expression was published more than 25 years ago [28]. Murine undifferentiated embryo cells treated with 5-azaeytidine, a potent inhibitor of DNA methylation, developed into various types of cells, including muscle and fat cells. These changes were inherited by the next generation of cells and it became obvious that
reducing DNA methylation reactivated certain genes, allowing the development of new cells from the original embryo. Today, two different pathways have been described for the inactivation of gene transcription by DNA methylation: 1) Methyl-CpGs can repel transcription factors directly by being present in the transcription factor binding sequence. Although regulation by such a mechanism \textit{in vivo} is relatively rare, some transcription factors, like Ets-1 [29] and the boundary element factor CTCF [30] are unable to bind DNA if the cytosines in their recognition sites are methylated. 2) DNA methylation can recruit proteins that bind methylated CpGs and subsequent inhibit transcription by remodelling the chromatin structure. This process is described in more detail below (page 18).

**Chromatin**

The fundamental subunit of chromatin is the nucleosome consisting of approximately 147 base pairs DNA, which is wrapped in two super-helical turns around an octamer of core histones; two of each of histones H2A, H2B, H3, and H4 [31]. The nucleosomes can be seen as “beads on a string” in an electron microscope under certain conditions, but are usually folded into higher order chromatin. (Figure 2) The chromatin is organized into domains of euchromatin and heterochromatin, which have different chromosomal architecture, transcriptional activity and replication timing. Euchromatin contains de-condensed, transcriptionally active regions, whereas heterochromatin is densely packed and contains mostly repetitive DNA but also some protein encoding genes. The structure and function of chromatin is highly dynamic, and regulation of gene expression often involves changes of its structure. Epigenetic mechanisms control gene expression by covalent modification of components within the chromatin or by remodelling the chromatin by ATP-dependent mechanisms.
Introduction

Figure 2. The higher order packing of DNA. The nucleosome is the lowest level of chromatin organization in which two super-helical turns of DNA are wrapped around a core of eight histones. The amino terminal tails of the histones protrude the nucleosome and are available for post-translational modification. The figure is modified after Felsenfeld and Groudine, 2003 [31]

Components of the main chromatin-modification and -remodelling complexes

Several components have a role in the chromatin modification and chromatin remodelling complexes. The main contributors are summarized in brief below:

*DNA methyltransferases*

DNA methyltransferases (DNMTs) are the enzymes responsible for both establishing and maintaining cellular DNA methylation. Today, three catalytically active DNMTs are known: DNMT1, DNMT3A, and DNMT3B. DNMT1 is ubiquitously expressed and has a significant preference for hemimethylated double-stranded DNA [32]. It is therefore commonly referred to
as the maintenance methyltransferase. DNMT3A and B on the other hand, initiate *de novo* methylation thereby establishing new methylation patterns [33] (figure 3). The expression of both DNMT3A and B are developmentally regulated.

**Figure 3. The role of DNA methyltransferases in CpG methylation.** DNMT3A and DNMT3B have the ability to fully methylate unmethylated DNA, which occur at the 5-position of cytosine in a CpG dinucleotide. This is called *de novo* methylation and targets both strands of the DNA double helix. Upon replication, the DNA becomes hemimethylated, as the newly synthesized strand is unmethylated whereas the template strand has kept its methylation. This “old” methylated DNA strand serves as a template for maintenance methylation by DNMT1, resulting in a fully methylated DNA duplex. The figure is modified after Reinhart and Chaillet, 2005 [34].

**Methyl-CpG-binding proteins**

Twelve methyl-CpG binding proteins have been identified in mammals, and among them MECP2, MBD1, MBD2, and MBD3 are the most studied [35]. Except for MBD3, they all
selectively bind methylated CpG sites [36]. These proteins can repress transcription by recruiting components that increases the packing of chromatin, thereby denying access to the transcriptional machinery. MECP2 for instance, represses gene expression by recruiting histone deacetylases as well as histone lysine methyltransferases [37]. In cancer, MBD proteins are associated with aberrantly methylated tumor suppressor genes and seem to constitute the mechanistic link between DNA methylation and gene silencing [38,39]. Interestingly, MBD4 has an alternative role as a DNA repair protein involved in removing T from 5mCpG-TpG mismatches [40], which counteracts the common 5-methylcytosine to thymine mutations. MBD4 has a mononucleotide repeat in the coding sequence and is frequently mutated in microsatellite unstable tumors [41]. Moreover, mice deficient of MBD4 have increased tumor formation [42].

Histone modification enzymes

Histones can store epigenetic information through a complex set of post-translational modifications [43]. The protruding amino terminal tails of the core histones can be acetylated, phosphorylated, methylated and/or ubiquitinylated, and the combined modification determines the gene activity [44]. Acetylation of histone lysines is distributed by histone acetyltransferases and is generally associated with transcriptional activation [45]. Several transcriptional co-activators, like p300/CBP have intrinsic histone acetyltransferase activity [46]. In contrast, histone deacetylases remove acetylation. The functional consequence of histone methylation depends on both the residue type (lysine or arginine) and specific site, e.g. methylation of lysine number four in histone H3 is linked to transcription [47], whereas methylation of lysine number nine in the same histone tail is associated with lack of transcription [48]. Several classes of histone methyltransferases have been identified [49]. Global histone modification patterns have recently been demonstrated to predict risk of prostate cancer recurrence [50].
Introduction

**ATP-dependent remodelling complexes**

The ATP-dependent chromatin remodelling enzymes can alter chromatin structure by disrupting or mobilizing nucleosomes in an energy-dependent manner. This increases the accessibility of nucleosomal DNA, which is a fundamental requirement for several steps in transcription. The three main classes of mammalian remodelling complexes – SWI2/SNF, Mi-2/NuRD, and ISW1 – contain different catalytic ATPase subunits and are associated with different additional proteins [51].

**Epigenetic regulation of gene expression – an interplay of multiple components**

Briefly reviewing the main contributors in epigenetic regulation of gene expression, the complexity seems striking. In this intricate regulation, DNA methylation plays a major, but not an independent role. In experimental systems, it seems that methylation at gene promoters does not lead to silenced transcription until additional proteins are recruited to the region, which mediates gene silencing [52]. Early models explaining the steps in this process were quite simplistic: the transcriptional repressor MECP2 binds methylated CpGs and recruits a complex containing a transcriptional co-repressor and a histone deacetylase. The subsequent deacetylation of histones would reduce gene transcription [53]. This model is still standing, but the continuous identification and functional investigation of components contributing in this process has led to a more complex story where the various proteins involved in chromatin modification and remodelling interact with one another to regulate chromatin structure and gene transcription. Figure 4 summarizes alternative links between DNA methylation, histone modification, and chromatin remodelling in eukaryotic cells.
Figure 4. Links between DNA methylation, histone modification and chromatin remodelling. DNA methylation, histone modification, and chromatin remodelling are involved in gene inactivation by silencing transcription. Three possible models of their interaction are presented. (A) DNA methylation directing histone methylation. \textit{De novo} methylated DNA recruits methyl-CpG-binding proteins (MBDs), which usually occur in a complex with histone deacetylase (HDAC). HDAC removes acetylation from the protruding histone tails leading to loss of gene transcription. The inactive state of chromatin can be stabilized by lysine 9 methylation of histone H3 by histone methyltransferases (HMTs). (B) Histone methylation directing DNA methylation. Methylated lysine 9 on histone H3 recruits heterochromatin protein 1 (HP1). HP1 recruits DNA methyltransferase through an unknown factor x, which maintains the DNA methylation and stabilizes the inactive chromatin. (C) Chromatin remodelling driving DNA methylation. Nucleosomal DNA can be unwound by chromatin remodelling complexes like ATRX and Lsh, which increases its accessibility to DNMTs, HDACs, and HMTs. The chromatin-remodelling protein (CRP) involved in the initial \textit{de novo} methylation has not yet been identified. Figure from En Li, 2002 [54]

DNA methylation plays a vital role in gene silencing, but whether it has a causative role in this process or merely is a consequence of inactivation is still debated. However, even if methylation changes do not arise first in epigenetic reprogramming, they are nevertheless important in preserving epigenetic states. Further, the unresolved question about causality does not undermine the promise of DNA methylation as a valuable biomarker (see discussion).
DNA methylation in normal development

The extent of DNA methylation changes in an ordered way during mammalian development. Immediately after fertilization both the parental and the maternal genomes are demethylated; the male genome by an active process that is completed within hours [55,56], and the maternal genome in a passive way during subsequent cleavage divisions [54]. After implantation, the genomes of the cells in the inner cell mass of the blastocyst have a high methylation level due to \textit{de novo} methylation, but the level tends to decrease in specific tissues during differentiation [23]. Mouse embryos homozygous for a mutation in the murine DNA methyltransferase gene have a recessive lethal phenotype and do not survive past mid-gestation [57], demonstrating that DNA methylation is essential for normal development.

The majority of CpG islands are unmethylated in normal tissue with the exception of a handful well-known cases: imprinted genes, x-chromosome genes in women, germ-line specific genes, and tissue specific genes. Intragenomic parasitic sequences and repetitive elements are additionally methylated in normal cells. Each of these categories will be discussed in brief in the following and are summarized in figure 5, along with the cancer specific changes of DNA methylation.
Genomic imprinting

The first imprinted gene identified was the mouse insulin-like growth factor 2 gene (Igf2) [59], and soon thereafter imprinting of the human homolog IGF2 was found [22]. Since then, several imprinted genes have been described and to date about 80 such genes have been isolated from the human and mouse genomes (see the Harwell Mouse Imprinting web site⁶ for a list of these genes). However, recent in silico analyses predicts that the murine genome contains as much as 600 potentially imprinted genes [60]. In contrast to the bulk of genes in the human genome, which are expressed from both the paternal and maternal allele, imprinted genes are

⁶ Web site: http://www.mgu.har.mrc.ac.uk/research/imprinting/
functionally haploid and have patterns of expression that depend on the parent of origin of their alleles.

The mechanism of genomic imprinting is complex and not completely understood, but it is established that DNA methylation plays a lead molecular role [61]. Around 80% of the imprinted genes are located in clusters, some of which are regionally regulated by imprinting centers or imprinting control elements. Imprinted genes are also unusually rich in CpG islands and the great majority have differences in DNA methylation between the parental alleles [62]. Deletions of these differentially methylated regions result in loss of imprinting [63]. The differentially methylated regions can have various properties: some are methylated in the inactive gene copy, whereas others are methylated in the active one. DNA methylation of the inactive allele is thought to induce a closed chromatin conformation with subsequent loss of gene expression [64]. The methylated regions in the active allele on the other hand, might contain silencers that are inactivated by methylation, perhaps by excluding repressor factors [65]. Epigenetic modification of boundary elements – DNA sequences that lie between two gene controlling elements preventing their interaction [62], as well as overlapping antisense transcripts [66], and post-transcriptional mechanisms [67] might also play a role in the regulation of imprinted gene expression.

Several of the genomic imprinted genes have roles in development and growth control, and imprinting might work to balance maternal and paternal demands on the rate of fetal growth. This was first suggested in a parent-offspring conflict model in 1989 [68]. Briefly described, during embryogenesis there is a conflict between the desire of the father to optimize the reproductive fitness of his offspring by promoting their growth, and the mother, who would benefit from conserving her own resources and distribute limited maternal resources equally to
all current and future offspring. This is exemplified by the paternally expressed fetal-specific growth factor \( \text{Igf2} \), which stimulates growth and the maternally expressed \( \text{Igf2r} \), which neutralizes \( \text{Igf2} \) by trafficking it into the lysosomes for degradation, thereby inhibiting embryonic growth [69]. However, not all imprinted genes have growth regulatory capabilities and are therefore difficult to fit into this model. An example is \( \text{Snrpn} \), which is imprinted in both humans and mice and thought to be an RNA splicing factor. Several alternative models explaining the evolution of genomic imprinting have been suggested [70-72], but the parent-offspring conflict model (also known as the kinship selection hypothesis) is still the most widely accepted.

Several human genetic diseases like the Beckwith-Wiedermann, Prader-Willi, and Angelman Syndrome have been associated with defects in genomic imprinting (reviewed in [73,74]). Among these, the Beckwith-Wiedermann Syndrome is linked to cancer as affected individuals are predisposed to embryonal tumors, the most frequent of which are Wilms’ tumor and adenocortical carcinoma [75] (see also “DNA methylation in cancer”).

**X-chromosome inactivation**

X-chromosome inactivation is a dose compensation mechanism to equalize the X-linked gene expression between men (one X chromosome) and women (two X chromosomes). Like imprinting, the hallmark of X-chromosome inactivation is monoallelic gene expression, and one of the female X chromosomes is therefore silenced early in the development. This inactivation is initiated by transcription of the \( \text{XIST} \) (X-inactive specific transcript) gene – a noncoding transcript that originates at the X inactivation centre [76]. The XIST RNA coats or “paints” one of the X chromosomes in \( \text{cis} \) [77] and triggers its inactivation by recruiting silencing complexes that induce chromatin changes. The resulting inactivated X chromosome is
very condensed and cytologically visible as a Barr body in the cell nucleus [78]. Once an X chromosome is inactivated in a cell, the inert state of the chromosome is clonally inherited during cell divisions.

The X chromosome inactivation in human epiblast cells of the inner cell mass occurs randomly with equal inactivation probabilities for the paternally and maternally inherited X [79]. Whereas the \textit{Xist} gene of the inactivated chromosome is transcriptionally active, the \textit{Xist} of the active chromosome is silenced by DNA methylation [80]. Several recently identified non-coding transcripts including \textit{Tsix}, \textit{DXPas34}, and \textit{Xce} are also transcribed from the X inactivation centre and might modulate the activity of the \textit{Xist} transcript (reviewed in [81]). The antisense \textit{Tsix} transcript, for example, initiated downstream of \textit{Xist}, acts as a negative regulator of \textit{Xist} expression [82].

DNA methylation is thought to be an important component of the inactivation of the X chromosome, but does not seem to be involved in its initial establishment as the \textit{Xist} activation and transcription precedes \textit{de novo} CpG island methylation in the mouse embryo [83]. However, maintenance of X inactivation requires DNA methylation in addition to other mechanisms, and several genes on the inactive chromosome are associated with hypermethylated CpG islands [83]. Furthermore, the epigenetic mechanisms that are involved in X inactivation and genomic imprinting seem to share several similarities and might have co-evolved in mammals [79].

**Tissue-specific genes**

DNA methylation has long been implied in tissue-specific gene expression control and for more than 25 years, this theory has been frequently cited [84,85]. Although DNA methylation
patterns are different from tissue to tissue, previous studies of CpG islands in tissue-specific
 gene promoters have failed to find a definite correlation with gene expression (e.g. [86]).
 Moreover, the gene expression of previous tissue-specific methylated candidates has been
 found to be unaffected by demethylation in methyltransferase-deficient mouse embryos [87].

Only recently has DNA methylation actually been proven to control tissue-specific gene
 expression, and \textit{SERPINB5} (maspin) was among the first genes to be identified [88].
 \textit{SERPINB5} expression is limited to certain types of epithelial cells, including those of the
 airway, breast, skin, prostate, and mouth. These cells harbor an unmethylated \textit{SERPINB5}
 promoter with acetylated histones and accessible chromatin structure. In contrast, non-
 expressing cells like skin fibroblasts, lymphocytes, bone marrow -, heart -, and kidney – cells
 have a fully methylated \textit{SERPINB5} promoter with hypo-acetylated histones and inaccessible
 chromatin [88]. Methylation of the first exon of \textit{DNAJD1 (MCJ)} has also been associated with
 loss of histone acetylation in a tissue-specific manner [89] and altogether, a handful tissue-
 specific genes controlled by DNA methylation have so far been identified (reviewed in [90]).

\textbf{Germ line specific genes}

Cancer/testis associated genes are specifically expressed in germ cells of the testis and
 occasionally in cells of female reproductive organs. The genes are not expressed in normal
 somatic tissues, but can be re-expressed in various types of human tumors [91]. \textit{MAGE1}
 (melanoma-associated antigen) was the first such gene to be identified [92], and today, about
 90 genes or isoforms have been described, including members of the MAGE, GAGE, PAGE,
 and XAGE families [93]. Except for their gene expression pattern and frequent X chromosome
 location, these genes have little in common. Their encoded proteins are thought to be important
 in the formation of mature functional spermatozoa, since different members are expressed
during various stages of spermatogenesis. However, the exact protein function has only been established for a small fraction of these gene products, and the remaining family members seem to encode both activators and repressors of proliferation and transcription, based on sequence homology with known proteins [93].

Detailed methylation analyses of cancer/testis associated genes have revealed that their CpG rich promoters are highly methylated in non-expressing somatic tissues, and mainly unmethylated in male germ cells [94]. DNA methylation is therefore the main - but not necessarily the only - regulator of the cancer/testis associated gene expression pattern. Does the up-regulated gene expression in various tumors make these genes oncogenes? In order to answer this, we need to establish whether they contribute to tumorigenesis or not [95]. However, the heterogeneous protein expression in human tumors might indicate that the cancer/testis associated genes do not have an initiating or crucial role in tumorigenesis [91].

**Intragenomic parasites and repetitive sequences**

DNA methylation protects mammary genomes from damage introduced by intragenomic parasites and repetitive sequences [96] (reviewed in ref [97]). Transposons, endogenous retroviruses, and other parasitic elements account for more than 35% of the human genome. Even though the majority are relics of once-active elements, and do no longer pose a threat to the genome, the remaining active elements can disturb the structure and deregulate gene expression in different ways. Several of these parasitic elements have the ability to move from one chromosomal location to another, in a process where the host genome is frequently mutated [98]. They can also cause illegitimate genomic rearrangements by their mere abundance, as the high frequency of homologous sequences of DNA increases the rate of non-allelic recombination. Finally, retrotransposons with strong constitutive promoters can produce
chimeric mRNAs representing a mix of retrotransposon sequences and endogenous exons [99]. Indeed, such chimeric retrotranscripts have been found in human tissues [100], but their functions need to be investigated further. In normal cells the parasitic elements are controlled by DNA methylation, as well as by direct transcriptional repression mediated by several host cell proteins. The methylation can inactivate the intragenomic parasites directly by promoter methylation, or indirectly by preventing the expression of additional genes required for transposition. Over time, spontaneous deamination of the methylated CpG sites leading to C to T transition mutations might destroy these elements.

**DNA methylation in cancer**

The epigenetic equilibrium described for the normal cell is dramatically disturbed during tumorigenesis. One of the first indications of this was seen in 1983 [22,90,101], when loss of DNA methylation at CpG dinucleotides was found in cancer genomes. Two decades, and more than 7000 articles later, we start to recognize the main processes by which DNA methylation is involved in cancer development (summarized in figure 5)

**Hypomethylation, genomic instability, gene activation, and loss of imprinting**

The genome of the cancer cell is hypomethylated in comparison with normal tissue, and the hypomethylated regions include repetitive DNA and intragenomic parasitic sequences [101]. The biological significance of this cancer-specific DNA hypomethylation has not been elucidated, but studies in mouse models confirm a causal connection between hypomethylation and tumor formation [102]. Mice with highly reduced levels of Dnmt1 and subsequent genomic hypomethylation, developed aggressive T cell lymphomas. Additionally, a high frequency of chromosome 15 trisomy was seen in these tumors, indicating that hypomethylation might lead to chromosomal instability [102]. A study of murine embryonic stem cells nullizygous for
*Dnmt1* and with increased chromosomal rearrangements corroborates this association [103]. Links between hypomethylation and genomic instability have also been shown in human primary tumors, including breast [104] and prostate [105], as well as in colon cancer cell lines [106]. Further insights into the possible relationship between hypomethylation and genomic instability can be gained from the only known human disease associated with a mutation in a DNA methyltransferase gene, the ICF syndrome. This syndrome is caused by an inactivating germ-line mutation in the *de novo* methylation *DNMT3B*, which results in loss of methylation at selected centromeric regions. In addition to immunodeficiency and facial abnormalities, patients with this syndrome display centromeric instability leading to profound chromosomal structural changes [107]. Although these patients do not develop cancer, similar regional losses of DNA methylation and chromosomal structural changes have been seen in many tumor types [108]. Altogether, the various reports mentioned above indicate that hypomethylation might lead to chromosomal instability in cancer. However, even though this is an attractive hypothesis, the full relationship between the two processes is not yet clear. Some reports, like the one from Lairds’ lab demonstrating fewer deletions in *Dnmt1*-deficient murine cells [109], present opposing results, leaving it up to future research to draw final conclusions.

Genomic hypomethylation has also been suggested to reactivate genes normally silenced by DNA methylation. Indeed, the cancer/testis associated genes mentioned above, inactivated by DNA methylation in normal somatic tissues, are re-expressed in several types of tumors [91]. The first hypomethylated oncogene, *HRAS*, was reported in 1983 [110]. Since then, an association between hypomethylation and over-expression have been demonstrated for a handful oncogenes, like *BCL2* in chronic lymphocytic leukemias [111], *ABCB1 (MDRI)* in myeloid leukemias [112], *S100A4* in colon adenocarcinomas cells [113], *SNCG (BCSG1)* in breast carcinomas and ovarian carcinomas [114], and *CCND2 (cyclin D2)* in gastric carcinoma...
Early reports also suggested *MYC* (*c-MYC*) as an oncogene activated by hypomethylation [116], however, subsequent publications revealed that the sites demethylated in cancer were not crucial for gene regulation [117]. Recently, several novel candidates have been added to the list of over-expressed hypomethylated genes by microarray analyses in pancreatic cancer: *CLDN4*, *LCN2*, *YWHAS* (*14-3-3σ*), *TFF2*, *MSLN*, and *PSCA* [118] and gastric cancer: *ELK1*, *RRAS*, *RHOB*, *RND1* (*RHO6*), and *MSX2* [119]. The increasing use of this type of large-scale technology, will improve the possibility of discovering more hypomethylated candidates in cancer. At present, the list of identified targets is overall short, especially compared with the opposite list of tumor suppressor genes inactivated by DNA hypermethylation. Further, it remains to be elucidated to what extent these activated oncogenes contribute to carcinogenesis.

Genomic hypomethylation can induce loss of imprinting, leading to either aberrant activation of the normally silent copy of a growth-promoting gene, or the silencing of the normally expressed allele of a growth-inhibiting gene. Because imprinted genes are functionally haploid and aberrations in only one allele is enough to disturb the activity, they increase our susceptibility to cancer. This is evident by the loss of imprinting of *IGF2* seen in Wilms tumor, leading to pathological biallelic expression [120]. Patients with Beckwith-Wiedemann syndrome, which have lost their *IGF2* imprinting either in the germline or early in development, have a 1000-fold higher incidence of this juvenile kidney tumor than the normal population [73]. Loss of *IGF2* imprinting is also frequently seen in colorectal cancer and even in normal colonic mucosa of about 30% of the same patients [121]. Only 10% of healthy individuals have loss of imprinting of *IGF2* [121], but this disrupted genomic imprinting appears to contribute to colorectal cancer risk [122]. Finally, several human tumor types are associated with aberrantly imprinted genes, other than *IGF2* [73].
**Hypermethylation and gene silencing**

In contrast to the questions that are raised regarding the exact function and consequence of hypomethylation in human cancer, the role of DNA hypermethylation is better understood. *De novo* methylation of CpG islands in gene promoters is associated with loss of expression, and examples of such inactivation can be found in as good as all cancer-related pathways (some of which are listed table 1) The first discovery of hypermethylation in a gene promoter in human cancer was that of the calcitonin gene in 1986 [123]. However, a link between CpG island promoter hypermethylation and transcriptional inactivation was not found until 1989, during the analyses of *RB* methylation [124]. Along with this discovery, the potential impact of inactivating DNA hypermethylation in tumorigenesis was realized [125].

The DNA hypermethylation in cancer seems to be a tissue specific event. Some genes are commonly methylated in a variety of tumors, like *p16INK4a* [126-128] and *RASSF1A* [129], whereas the majority are highly specific with respect to the tissue of tumor origin [130]. One example is *GSTP1*, which is hypermethylated in the majority of tumors in the prostate [131], liver, and to a less extent in breast [132] but is largely unmethylated in other cancer types [132]. The mechanism for this discrimination of tissue origin remains to be determined, but might be connected to the biology of the separate tumor types. In accordance with the Darwinian perspective of tumor development, we might assume that only genes conferring a selective advantage to the specific tumor type will be inactivated.
Introduction

Table 1. A selection of cellular pathways and processes affected by aberrant DNA methylation in cancer.

<table>
<thead>
<tr>
<th>Pathways</th>
<th>Genes</th>
<th>Selected references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered cell cycle control</td>
<td>(RB1), (p16^{INK4a}), (p14^{ARF}), (CDKN2B) ((p15)), (CDKN1A) ((p21)), (TP73)</td>
<td>[128,134-138]</td>
</tr>
<tr>
<td>Repair of DNA damage</td>
<td>(MLH1), (MGMT), (BRCA1)</td>
<td>[139-141]</td>
</tr>
<tr>
<td>The TP53 network</td>
<td>(p14^{ARF}), (TP73)</td>
<td>[135,138]</td>
</tr>
<tr>
<td>The WNT signaling pathway</td>
<td>(APC), (CDH1), (SFRPs)</td>
<td>[142-144]</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>(DAPK1), (CASP8), (PYCARD) ((TMS1))</td>
<td>[145-147]</td>
</tr>
<tr>
<td>Tumor cell invasion or tumor architecture</td>
<td>(CDH13), (VHL), (STK11), (TIMP3), (THBS1)</td>
<td>[132,148-151]</td>
</tr>
<tr>
<td>Hormonal response</td>
<td>(ER) ((ESR1)), (PGR) ((PR)), (AR), (RAR) ((RAR)), (RB) ((RBP))</td>
<td>[152-156]</td>
</tr>
<tr>
<td>Cytokine signaling</td>
<td>(SOCS1), (SOCS3)</td>
<td>[157,158]</td>
</tr>
</tbody>
</table>

DNA methylation, ageing and diet

Environmental factors can induce epigenetic changes and thereby contribute to the development of abnormal phenotypes [159]. The amount and pattern of DNA methylation in somatic cells have been associated with increased age; a progressive loss of overall methylation is seen in aging rodents [160,161], as well as during \textit{in vitro} culturing of human fibroblasts [162]. Distinct from this global decrease in DNA methylation, specific genes have been reported to be hypermethylated in tissues of aging individuals. The estrogen receptor (\(ER\), \(ESR1\)) was the first gene in which methylation was demonstrated to increase as a direct function of age in normal colon mucosa [163]. \(IGF2\) and \(MYOD1\) are additional examples of unmethylated genes in young individuals that become progressively methylated with age in normal tissues [164]. Increased age is thought to be one of the greatest risk factors for cancer, and age-related methylation might be a fundamental predisposing event to the neoplastic transformation [165]. However, it is important to keep in mind that the methylation status of the majority of genes examined seems to be intact during ageing.
Loss of DNA methylation has also been linked to nutrition, as lack of cellular S-adenosyl-methionine, has been shown to predispose to cancer [166]. The DNA methyltransferases are dependent on a correct supply of this universal methyl donor in order to exercise their function and methylate DNA. Rodents fed a methyl-deficient diet, which diminishes tissue concentrations of S-adenosyl-methionine, display genomic DNA hypomethylation in the liver [167] and develop hepatocellular carcinomas [168]. In general, reduced amounts of folate - a dietary supplement important in generation of methyl groups - have been associated with genomic instability [169,170] and neural tube defects [171] in addition to genomic hypomethylation [172]. Several studies have reported an inverse association between folate intake and colorectal cancer incidence (reviewed in [173]). Moreover, polymorphisms in the methylene-tetrahydrofolate reductase (MTHFR), which is involved in folate metabolism, also affects the risk for colorectal cancer [173].

**DNA methylation and mutations**

The reversible epigenetic aberrations seen in cancer can induce several known irreversible genetic changes. In addition to the microsatellite unstable phenotype seen in MLH1 deficient tumors, cells harboring MGMT hypermethylation frequently contain G:C to A:T transition mutations in their DNA sequence. Active MGMT removes alkylation at the O6-position of guanine, which per se can base pair with thymine in addition to cytosine during replication and introduce stable mutations [174]. G:C to A:T mutation in both TP53 and KRAS2 have been associated with MGMT inactivation [175,176]. Moreover, the modified cytosine base, 5-methylcytosine, can also influence tumorigenicity by causing C-T transitions. Methylated CpG dinucleotides are in fact the single most important mutational target in the tumor suppressor
TP53 [177]. In general, 5-methylcytosines are responsible for more than 30% of all known disease-related point mutations [178].

**Testicular germ cell tumors – a curable rare malignancy arising in young men**

Testicular germ cell tumor (TGCT) is the most common malignancy in young males and comprises about 98% of all testicular neoplasms [179]. Three different entities of TGCT can be distinguished, teratomas and yolk sac tumors of newborn and infants, seminomatous and nonseminomatous germ cell tumors of adolescents and young adults, and spermatocytic seminoma of elderly men. Tumors from these three groups originate from germ cells at different stages of development and are epidemiologically, clinically, and histologically diverse [180]. The entity of seminomas and nonseminomas of young men is by far the most common form of TGCT, and is also the focus of this thesis. Only about 1-5% of all testicular cancers are familial cases, where brothers and sons to testicular cancer patients have a several-fold increased risk of developing a testicular tumor compared with the general population [181-184]. From linkage analyses in families with two or more cases of testicular cancer, chromosome arms 3q, 5q, 12q, 18q, and Xq have been suggested to contain testicular cancer susceptibility loci. However, besides from the testicular germ cell tumor gene 1 (TGCT1) residing in the Xq27 locus, relevant susceptibility genes remain unknown [185,186].

**Incidence, treatment, and outcome**

The incidence of TGCTs varies with geographic location and race, and the Scandinavian countries rank high on the list [180,187,188]. The current age-adjusted incidence rate of testicular cancer in Norway is approximately 11 per 100,000 males per year (figure 6). Patients
diagnosed with TGCT have their testis surgically removed (orchiectomy) and subsequent treatment, including radiotherapy and chemotherapy, is administered to the patient based on the histology and clinical stage of the tumor. Today, virtually all patients with localized and regional disease survive, representing the highest survival rate for any malignancy in men. However, until the introduction of the efficient cisplatin-based chemotherapy in the late 1970s, the overall survival rate for patients diagnosed with TGCT was low. TGCTs can metastasize to peripheral sites and patients with metastases at time of diagnosis receive chemotherapy (a combination of bleomycin, etoposid, and cisplatin) followed by resection of residual disease. Approximately 80-90% of patients with metastatic TGCT survive beyond five years after diagnosis.

Figure 6. Incidence rate for testicular cancer compared to that of other cancers in Norway. Left panel (A), trends in age-standardized incidence rates. Right panel (B), age-specific incidences, 1997-2001. The raw data were obtained from the Norwegian Cancer Registry; http://www.kreftregisteret.no

7 The Norwegian Cancer Registry’s web site: http://www.kreftregisteret.no/
Histopathology

The TGCTs are divided into two main histological classes, seminomas and nonseminomas [192,193], both suggested to develop from the precursor stage intratubular germ cell neoplasia (ITGCN; also named carcinoma in situ) [194]. ITGCN is initiated during fetal life from a primordial germ cell (PGC) or a gonocyte [194,195] and is seen adjacent to invasive germ cell tumors in the majority of cases [196,197]. The undifferentiated seminomas morphologically resemble the intratubular germ cell neoplasia, whereas the nonseminomas include several histological subtypes along complete differentiation lineages. Cells of the undifferentiated pluripotent embryonal carcinoma may differentiate into teratomas, containing tissues from all three germ layers, or they may differentiate along extra-embryonal lineages into yolk sac tumors or choriocarcinomas [192,194,198].

Genetic changes

TGCTs are characterized by excess genetic material of the short arm of chromosome 12 [199,200]. This is usually due to the presence of isochromosome 12p (i(12p)), detected in about 80% of the TGCTs, and initially described by Atkin and Baker in 1982 [201]. Close to all of the remaining tumors display gains or amplifications of 12p sequences [202], but only few reports have found i(12p) in ITGCN [203-205]. TGCTs are typically aneuploid. Whereas ITGCN and seminomas usually have chromosome numbers in the hyper-triploid range [206], the nonseminomas are consistently hypo-triploid [207,208]. Since aneuploidy and gain of 12p sequences is found in virtually all TGCTs, they are both considered to be early changes in the tumorigenesis. However, polyploidization is suggested to precede the formation of i(12p) as the presence of this isochromosome in ITGCN is still debated [180]. A recent study analyzing micro-dissected ITGCN samples suggests that the gain of extra chromosome 12 material most
likely is associated with a more malignant progression of the cells, rather than an early event in the neoplastic transformation [204]. In addition to i(12p), several recurrent chromosomal changes as well as many molecular changes have been reported in TGCT [193,209-212]. Net loss is commonly seen from chromosomes 4, 5, 11, 13, and 18, whereas chromosomes 7, 8, 12, 17, and X typically show gains in TGCT [213]. Although many of the target genes for the chromosomal aberrations remain unknown, several genes are found altered in these tumors, such as FHIT, JUP, and GRB7 [214-216]. Moreover, expression profiling of TGCT have identified several genes with significant altered expression in tumors compared with normal testis tissue [215,217-220].

**Epigenetic changes**

Like other cancer types, the genome of the TGCTs is hypomethylated. However, the seminomatous TGCTs are more extensively hypomethylated than are the nonseminomas [221]. Significant epigenetic differences between these two major histological subgroups can also be seen from DNA promoter methylation analyses [221,222]. Nonseminomas show frequent CpG island hypermethylation, comparable to other solid tumors, whereas seminomas have almost no methylation [221]. Studies of individual target genes in testicular tumorigenesis confirm this (table 2). Moreover, the separate nonseminomatous histological subgroups show variable frequencies of hypermethylation. Teratomas and yolk sac tumors seem to be the two most commonly methylated subtypes [223,224]. Since TGCTs arise from primordial germ cells (PGCs) at a stage in the development where they undergo epigenetic reprogramming [225,226], the DNA methylation in TGCTs should be evaluated in this context.
TGCTs consistently express both parental alleles of genes that are normally imprinted [227-229], which reflects that the parental imprinting has been erased in the cells they originate from.

Several studies have indicated an involvement of the X chromosome in testicular germ cell tumorigenesis [230-232]. Indeed, a numerical increase in X chromosomes is commonly observed in these tumors [213,230] and might explain the expression of the \textit{XIST} gene reported in TGCTs, which is usually only expressed in females [233,234]. The transcription of the \textit{XIST} gene initiates the process of X chromosome inactivation, however, in TGCTs the multiple X chromosomes seem to be predominantly hypomethylated and active in spite of the \textit{XIST} expression [234]. Nonetheless, the methylation status of \textit{XIST} in male-derived plasma has been suggested as a tumor marker for testicular cancer [235], but its usefulness is still debated [236].
<table>
<thead>
<tr>
<th>Gene</th>
<th>Methylation %</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All TGCTs</td>
<td>Nonseminomas</td>
</tr>
<tr>
<td>RASSF1A</td>
<td>22-71</td>
<td>29-83</td>
</tr>
<tr>
<td>MGMT</td>
<td>21-46</td>
<td>20-69</td>
</tr>
<tr>
<td>SCGB3A1</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td>HIC1</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>BRCA1</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>HOXA9</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>APC</td>
<td>10</td>
<td>24-29</td>
</tr>
<tr>
<td>FHIT</td>
<td>7</td>
<td>6-29</td>
</tr>
<tr>
<td>HOXB5</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>CDH13</td>
<td>9-12</td>
<td>0</td>
</tr>
<tr>
<td>CDH1</td>
<td>4-11</td>
<td>0</td>
</tr>
<tr>
<td>RARB</td>
<td>8</td>
<td>5-6</td>
</tr>
<tr>
<td>FANCF</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>EMX2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>MSX1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>RUNX3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>SORBS1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>XPA</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>hMLHL1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>TIMP3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>NKX3.1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>NME2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GSTP1</td>
<td>0-1</td>
<td></td>
</tr>
<tr>
<td>APAF1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BTG1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CGGBP1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CGRRF1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DAPK</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DLX6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LDHA</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MSX2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NEDD1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NME1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>p14ARF</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>p16INK4a</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RB1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SMARCC2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TP73</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Genes analyzed for promoter hypermethylation in TGCTs.** The genes listed in the table are sorted by hypermethylation frequencies, which are given in percentages for all TGCT (including both seminomas and nonseminomas) and for nonseminomas and seminomas separately (when this is reported). The germ line specific genes (e.g. MAGE) are not included in this table. Note: ref [238] and [223] come from the same research group and might include overlapping samples. Ref [241] is based on a single sample, a mediastinal germ cell tumor. Ref [242] is excluded since all nine genes analyzed in TGCT were unmethylated.
Colorectal cancer – a common malignancy with poor outcome

Up to 25% of all colorectal cancer cases can be associated with a positive family history of the disease [243], and among these approximately 5% are due to known hereditary syndromes [244]. Hereditary nonpolyposis colorectal cancer (HNPCC) is the most common, but several other genetic syndromes causing inherited predisposition for colorectal cancer have been described, including familial adenomatous polyposis (FAP) and MYH polyposis caused by germline mutations in the \( APC \) and \( MYH \) genes, respectively [243]. However, the vast majority of all colorectal cancer cases arise sporadically. This latter group is the focus of the present thesis.

Incidence, treatment, and outcome

Colorectal cancer is a common disease in the Western countries [245] and arises with comparable frequencies in men and women. The age-adjusted incidence rate of colon cancer in Norway is approximately 25 per 100,000 people per year (Figure 7), and the total number of new cases with colon or rectal carcinoma was in 2002 reported to be 3291\(^8\). Surgery is the main and most efficient treatment for colorectal cancer and all patients with primary tumors and local recurrence will have their tumor surgically removed. Patients with advanced local, regional, or metastatic disease, require additional therapy, like chemotherapy and radiation. In Norway, 5-fluorouracil (5-FU) in combination with calciumfolinate (leukovirin) is the most common chemotherapy for this patient group. The prognosis of colorectal cancer patients depends on the stage of the tumor at diagnosis. Overall, the survival is poor, and approximately 45% of the patients die within five years\(^8\).

\(^8\) The Norwegian Cancer Registry’s web site: http://www.kreftforeningen.no
Introduction

Figure 7. Incidence rate for colon cancer compared to that of other cancers in Norway. The raw data were obtained from the Norwegian Cancer Registry; http://www.kreftregisteret.no

Histopathology

Distinct histopathologic as well as molecular differences are not only seen among the carcinomas, but also among the polyps. Several hyperplastic and dysplastic benign lesions have so far been described in the colorectum, including *e.g.* aberrant crypt foci (ACF), the earliest morphological identifiable precursor of epithelial neoplasia, hyperplastic polyps, hyperplastic polyps variants, mixed polyps/serrated adenomas, and traditional adenomas [246-249]. Most colorectal carcinomas seem to arise from benign adenomas, which gradually progress through increases in size, dysplasia, and villous complexity [250]. Hyperplastic polyps have long been considered to have no malignant potential. However, several recent molecular studies have indicated that these polyps may be neoplastic rather than true hyperplastic [251,252].

The adenoma-carcinoma sequence

Based on histopathological findings, Muto and co-workers suggested in 1975 that cancers of the colon and rectum have evolved through a polyp-cancer sequence [253]. Fearon and Vogelstein later proposed a model for the genetic basis of this development [250], which also included the hypomethylation seen in very small adenomas [254]. In contrast to many other
tumor types, the pre-invasive lesions, as well as malignant and metastatic tumors can be isolated from the patient, which makes this cancer type a good model for tumor developmental studies. Both the initiation and progression of sporadic colorectal carcinomas are characterized by the gradual accumulation of genetic and epigenetic alterations and figure 8 summarizes the main changes in this process.

Figure 8. Genetic and epigenetic changes in the adenoma - carcinoma sequence seen in left- versus right-sided carcinogenesis. Tumors with microsatellite instability (MSI) are associated with right-sided location and a diploid genome. Tumors with chromosome instability (CIN) are associated with left-sided and rectum-location and an aneuploid genome. Main epigenetic changes are marked in green and discussed in detail in the text. Genetic changes associated with MSI tumors are marked in red whereas genetic changes associated with CIN tumors are marked in blue. Serrated adenomas and hyperplastic polyps (not illustrated here) can also give rise to colorectal carcinomas.

**Genomic instability**

Loss of genomic stability has been proposed to be a key in cancer formation [255]. Two distinct molecular pathways leading to genomic instability have been described for colorectal
Introduction
cancer, the so-called chromosome instability (CIN) - and microsatellite instability (MSI) - pathways. Tumors with the CIN phenotype are characterized by multiple chromosomal aberrations [256]. The cause(s) responsible for the instability process in these tumors remain mostly unknown, although many of the genes encoding proteins involved in the mitotic-spindle - and the DNA replication - checkpoints have been closely analyzed for changes [18]. In contrast, the molecular mechanism causing MSI has been well characterized. The mismatch repair system recognizes and repairs base-pair mismatches that occur during DNA replication. If components in this complex are inactivated, multiple insertions and deletions of short nucleotide repeats will be seen in coding as well as in non-coding sequences throughout the genome [18]. Approximately 15% of all sporadic colorectal carcinomas display MSI [257]. The CIN and MSI tumors are clinicopathologically different. The CIN tumors are aneuploid, whereas the MSI tumors are usually diploid [258]. Furthermore, the CIN tumors are often found in the left side of the colon, whereas MSI tumors are associated with right-sided location [259,260]. Patients with MSI tumors have a better survival rate than have CIN patients [260,261]. Moreover, MSI cancers are more likely to be present at a more advanced age [262] and to be associated with the female gender [263].

Epigenetic changes

The time of hypomethylation in neoplastic transformation can vary from tumor type to tumor type. In most cancers, the hypomethylation is an early event, but exceptions have been demonstrated. In adenocarcinomas of the prostate hypomethylation is observed only in a fraction of the early cases, whereas extensive hypomethylation occurs mostly at the metastatic stage [105,264-266]. In colorectal tumorigenesis, the hypomethylation occurs at an early stage. Even very small adenomas reveal loss of methyl groups in the DNA [254,267]. Simultaneous
regional DNA hypermethylation has been identified in these lesions, leading to transcriptional inactivation of the mismatch repair gene *MLH1* among others [268]. Hypermethylation of *MLH1* was first reported in 1997 [17,139,269] and today, approximately 80% of the sporadically occurring MSI tumors are estimated to be caused by this epigenetic inactivation [270]. *APC*, initially identified as a common gene target for mutations in the adenoma-carcinoma sequence, can also be inactivated by DNA hypermethylation. However, the frequency of *APC* methylation is significantly lower than that of *MLH1* in sporadic tumors [271]. Genes reported to be inactivated by DNA hypermethylation in colorectal carcinomas are listed in table 3. Some of these genes are hypermethylated in adenomas as well as in carcinomas, and may thus represent early and even initiating changes in the tumorigenesis. Indeed, Wynter and co-workers have suggested that a subgroup of colorectal carcinomas (MSI-low/MSS) develop from hyperplastic polyp variants by a pathway including hypermethylation and subsequent inactivation of *MGMT* combined with mutation of *KRAS2* [249]. The methylation frequency of *MGMT* is similar in colorectal adenomas and carcinomas, further supporting a role for this gene in tumor initiation [272].
<table>
<thead>
<tr>
<th>Genes</th>
<th>Methylation</th>
<th>Comments</th>
<th>Selected References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Promoters methylated in colorectal tumorigenesis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>Medium</td>
<td></td>
<td>[273]</td>
</tr>
<tr>
<td>ADAMTS1</td>
<td>High</td>
<td></td>
<td>[274]</td>
</tr>
<tr>
<td>APC</td>
<td>Low</td>
<td></td>
<td>[142]</td>
</tr>
<tr>
<td>BNIP3</td>
<td>High</td>
<td></td>
<td>[275]</td>
</tr>
<tr>
<td>CD44</td>
<td></td>
<td></td>
<td>[276]</td>
</tr>
<tr>
<td>CDH1</td>
<td>Medium</td>
<td></td>
<td>[272]</td>
</tr>
<tr>
<td>CDH13</td>
<td>Medium</td>
<td></td>
<td>[277]</td>
</tr>
<tr>
<td>CDH4</td>
<td>High</td>
<td></td>
<td>[278]</td>
</tr>
<tr>
<td>CDKN2B</td>
<td>Low-high</td>
<td>Disagreement regarding methylation frequency</td>
<td>[279], [132]</td>
</tr>
<tr>
<td>CDX1</td>
<td>High</td>
<td></td>
<td>[279]</td>
</tr>
<tr>
<td>CDX2</td>
<td>Medium</td>
<td></td>
<td>[280]</td>
</tr>
<tr>
<td>COL1A2</td>
<td></td>
<td></td>
<td>[281]</td>
</tr>
<tr>
<td>CRABP1</td>
<td>Medium</td>
<td></td>
<td>[274]</td>
</tr>
<tr>
<td>DAB2IP</td>
<td>Medium</td>
<td></td>
<td>[282]</td>
</tr>
<tr>
<td>DAPK1</td>
<td>Low</td>
<td></td>
<td>[283]</td>
</tr>
<tr>
<td>EPHA7</td>
<td>Medium</td>
<td></td>
<td>[284]</td>
</tr>
<tr>
<td>FAS</td>
<td>Low-medium</td>
<td>Disagreement regarding methylation frequency</td>
<td>[285], [286]</td>
</tr>
<tr>
<td>FHIT</td>
<td>Medium</td>
<td></td>
<td>[287]</td>
</tr>
<tr>
<td>GATA4</td>
<td></td>
<td></td>
<td>[288]</td>
</tr>
<tr>
<td>GATA5</td>
<td></td>
<td></td>
<td>[288]</td>
</tr>
<tr>
<td>GSTP1</td>
<td>Low</td>
<td></td>
<td>[272]</td>
</tr>
<tr>
<td>HIC1</td>
<td>High</td>
<td></td>
<td>[289]</td>
</tr>
<tr>
<td>HOXA9</td>
<td>Medium</td>
<td></td>
<td>[290]</td>
</tr>
<tr>
<td>HRK</td>
<td></td>
<td></td>
<td>[291]</td>
</tr>
<tr>
<td>HS3ST2</td>
<td>High</td>
<td>Alias: 3OST2</td>
<td>[292]</td>
</tr>
<tr>
<td>HLTF</td>
<td>Medium</td>
<td></td>
<td>[293]</td>
</tr>
<tr>
<td>ID4</td>
<td>Medium</td>
<td></td>
<td>[294]</td>
</tr>
<tr>
<td>LPHN2</td>
<td>Medium</td>
<td>Alias: KIAA0786</td>
<td>[295]</td>
</tr>
<tr>
<td>MGMT</td>
<td>Medium</td>
<td></td>
<td>[140]</td>
</tr>
<tr>
<td>MINT1,2,12, 17, 25, 27, 31*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLH1</td>
<td>High</td>
<td></td>
<td>[139]</td>
</tr>
<tr>
<td>NR3C1</td>
<td>Medium</td>
<td></td>
<td>[274]</td>
</tr>
<tr>
<td>p14ARF</td>
<td>Medium</td>
<td></td>
<td>[297]</td>
</tr>
<tr>
<td>p16INK4a</td>
<td>Medium</td>
<td></td>
<td>[297]</td>
</tr>
<tr>
<td>TP73</td>
<td>High</td>
<td></td>
<td>[279]</td>
</tr>
<tr>
<td>PIK3CG</td>
<td>Medium</td>
<td></td>
<td>[298]</td>
</tr>
<tr>
<td>PRDM2</td>
<td>Low</td>
<td>Alias: RIZ, RIZ1</td>
<td>[299]</td>
</tr>
<tr>
<td>PTEN</td>
<td>Low-high</td>
<td>Disagreement regarding primer design</td>
<td>[300], [287]</td>
</tr>
<tr>
<td>PTGIS</td>
<td>Medium</td>
<td></td>
<td>[301]</td>
</tr>
<tr>
<td>PTGS2</td>
<td>Low</td>
<td>Alias: COX2</td>
<td>[302]</td>
</tr>
<tr>
<td>PTPRO</td>
<td></td>
<td></td>
<td>[303]</td>
</tr>
<tr>
<td>PYCARD</td>
<td>Medium</td>
<td>Alias: TMS1, ASC</td>
<td>[304]</td>
</tr>
<tr>
<td>RAB32</td>
<td></td>
<td></td>
<td>[303]</td>
</tr>
<tr>
<td>RARB</td>
<td></td>
<td></td>
<td>[305]</td>
</tr>
<tr>
<td>RASSF1A</td>
<td>Low</td>
<td></td>
<td>[306]</td>
</tr>
<tr>
<td>RASSF2A</td>
<td>High</td>
<td></td>
<td>[307]</td>
</tr>
<tr>
<td>RBP1</td>
<td>Medium</td>
<td></td>
<td>[308]</td>
</tr>
</tbody>
</table>
Introduction

<table>
<thead>
<tr>
<th>Gene</th>
<th>Methylation Level</th>
<th>Alias</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUNX3</td>
<td>Medium</td>
<td></td>
<td>[309]</td>
</tr>
<tr>
<td>SCGB3A1</td>
<td>Low</td>
<td>Alias: HIN-1, HIN1</td>
<td>[310]</td>
</tr>
<tr>
<td>SEZ6L</td>
<td></td>
<td></td>
<td>[295]</td>
</tr>
<tr>
<td>SFRP1</td>
<td>High</td>
<td></td>
<td>[295]</td>
</tr>
<tr>
<td>SFRP2</td>
<td>High</td>
<td></td>
<td>[295]</td>
</tr>
<tr>
<td>SFRP4</td>
<td>Medium</td>
<td></td>
<td>[295]</td>
</tr>
<tr>
<td>SFRP5</td>
<td>Medium</td>
<td></td>
<td>[295]</td>
</tr>
<tr>
<td>SLC5A8</td>
<td>Medium</td>
<td></td>
<td>[311]</td>
</tr>
<tr>
<td>SLIT2</td>
<td>High</td>
<td></td>
<td>[312]</td>
</tr>
<tr>
<td>SOCS1</td>
<td>Low</td>
<td></td>
<td>[313]</td>
</tr>
<tr>
<td>STK11</td>
<td>Low</td>
<td>Alias: LKB1</td>
<td>[150]</td>
</tr>
<tr>
<td>THBS1</td>
<td>Low</td>
<td></td>
<td>[272]</td>
</tr>
<tr>
<td>TIMP3</td>
<td>Low</td>
<td></td>
<td>[272]</td>
</tr>
<tr>
<td>TMEFF2</td>
<td>High</td>
<td>Alias: HPP1, TPEF</td>
<td>[314]</td>
</tr>
<tr>
<td>WIF1</td>
<td>High</td>
<td></td>
<td>[315]</td>
</tr>
<tr>
<td>WT1</td>
<td>Medium</td>
<td></td>
<td>[279]</td>
</tr>
</tbody>
</table>

Promoters methylated in ageing large bowel and in colorectal tumorigenesis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Methylation Level</th>
<th>Localization</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CXX1</td>
<td>Medium</td>
<td>X Chromosome localization</td>
<td>[279]</td>
</tr>
<tr>
<td>ER (ESR1)</td>
<td>High</td>
<td>Age-specific</td>
<td>[163]</td>
</tr>
<tr>
<td>IGF2</td>
<td>Medium</td>
<td>Imprinted and age-specific</td>
<td>[316]</td>
</tr>
<tr>
<td>MEST</td>
<td>Medium</td>
<td>Imprinted and age-specific,</td>
<td>[317]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>alias: PEG1</td>
<td></td>
</tr>
<tr>
<td>MINT4, 6, 24, 32*</td>
<td>High</td>
<td>Age-specific</td>
<td>[296]</td>
</tr>
<tr>
<td>MYOD1</td>
<td>High</td>
<td>Age-specific</td>
<td>[279]</td>
</tr>
<tr>
<td>TUSC3 (N33)</td>
<td>High</td>
<td>Age-specific</td>
<td>[318]</td>
</tr>
</tbody>
</table>

Table 3. Genes methylated in colorectal carcinomas. A summary of published data on methylation in normal and cancerous colon epithelium. The genes are listed in alphabetical order. Methylation is categorized into three groups according to the prevalence: low, range 1-20%; medium, range 21-60%; high, range 61-100%. For brevity, only one reference is listed per gene, except when there are major disagreements in the published literature. *MINT is not a gene symbol, but a term for clones “Methylated In Tumor” identified by methylated CpG amplification (MCA) followed by representational difference analysis.

CpG island hypermethylation is more frequent among MSI-positive carcinomas with proximal location than in left-sided MSS tumors [319-322]. The same association is found for adenomas [272], indicating that this type of epigenetic regulation plays a more important role in proximal colon tumorigenesis than in distal colon tumorigenesis.
CIMP or CIMPly not?

In 1999, Toyota and co-workers analyzed a panel of differentially methylated DNA sequences in colorectal tumors [296]. Two distinct groups of tumors were identified, a group with frequent methylation of three or more of the seven loci analyzed and a group with extremely rare methylation. Based on the concordant hypermethylation of multiple CpG islands in the first group, the authors suggested a third pathway to colorectal tumorigenesis, the CpG island methylator phenotype (CIMP) [296,319]. The CIMP concept has since been extensively discussed. Some studies support the original findings [299,323] whereas other do not [272,324]. The controversy stems from the fact that only a subset of CpG islands appears to be affected. A study from Peruchos’ group, which included a broader analysis of CpG island hypermethylation than the original paper, failed to confirm the existence of CIMP, as the hypermethylation distribution was continuous rather than bimodal [321]. The main problem with CIMP at the moment therefore seems to be its lack of a consensus definition.

However, the CIMP concept is extremely interesting in a cancer perspective. If this third pathway for colorectal tumorigenesis holds true, it will have several important implications. First of all, methylation in cancer would have a traceable cause. Analyses of genes potentially involved in DNA methylation in patients with familial clustering of CIMP or DNA methylation could identify the cause for this defect in epigenetic control. Furthermore, early detection of CIMP and subsequent epigenetic intervention could specifically prevent the formation of CIMP-positive cancers [325].
AIMS OF THE PROJECT

The aims of this project are based on the recognition of aberrant DNA methylation as a common mechanism to alter gene expression in human cancers and that such changes represent promising biomarkers in cancer diagnostics as well as in disease monitoring.

We aimed to identify novel epigenetically regulated target genes contributing to the development of cancer disease in the testis and in the large bowel, by combining the candidate gene approach with global genomic studies.

Further, we aimed to identify epigenetic differences in the two major histological subclasses of TGCT, seminomas and nonseminomas, as well as along the embryonal and extra-embryonal differentiation lineages of the latter group.

Finally, we aimed to gain novel insights into the epigenetics paralleling the tumorigenesis in the large bowel, by comparing methylation profiles of benign and malignant tumors stratified according to known molecular, pathological, and clinical data.
**RESULTS IN BRIEF**

**Paper I.** “Frequent promoter hypermethylation of the O\(^\text{6}\)-Methylguanine-DNA Methyltransferase (MGMT) gene in testicular cancer” In this study we addressed the potential involvement of MGMT and CDKN2A in primary testicular germ cell tumors. The DNA methylation status of the respective gene promoters, and the relative amount of parental alleles at selected loci in the chromosomal map position of the genes were investigated in a series of primary tumors. We found that MGMT was hypermethylated in 46% of testicular germ cell tumors, and that the hypermethylation was more common in nonseminomas 24/35 (69%), than in seminomas 8/33 (24%), \((P < 0.001)\). The MGMT expression was analyzed in 20 primary tumor samples by immunohistochemistry and six of seven methylated samples had lost all protein expression. CDKN2A on the other hand, was unmethylated in all samples analyzed. Allelic imbalances were frequently seen in eight microsatellite markers within and flanking MGMT and CDKN2A at chromosome bands 10q26 (50/70; 71%) and 9p21-22 (47/70; 67%), respectively. By a genome-wide RLGS\(^9\) study of CpG islands in testicular germ cell tumors, Smiraglia and co-workers showed in 2002 that hypermethylation is significantly more frequent in nonseminomas than in seminomas. The present study confirmed this in a large panel of primary tumors including all histological subtypes. Moreover, our data were the first to show that epigenetic inactivation of MGMT contributes to the development of nonseminomatous testicular cancer.

**Paper II.** “The loss of NKX3.1 expression in testicular- and prostate-cancers is not caused by promoter hypermethylation” In a previous report we have demonstrated that the homeodomain-containing transcription factor NKX3.1 protein is frequently lost in TGCT.

\(^9\) See Appendix I for complete list of abbreviations
This has also been shown in some studies of prostate adenocarcinomas. Since no gene mutations have been identified in the *NKKX3.1* gene, epigenetic changes have been hypothesized to be responsible for the reduced expression. We therefore analyzed the methylation status of the *NKKX3.1* promoter in primary TGCT (n = 55), intratubular germ cell neoplasias (n = 7), germ cell tumor cell lines (n = 3), primary prostate adenocarcinomas (n = 20), and prostate cancer cell lines (n = 3) by MSP and bisulphite sequencing. Only one sample (a TGCT) harbored hypermethylation of *NKKX3.1*. In conclusion, down-regulation of *NKKX3.1* expression is generally not caused by gene promoter hypermethylation. However, other epigenetic mechanisms responsible for the reduced gene expression cannot be excluded. Noteworthy, during the bisulphite sequencing we identified a possible polymorphism in the *NKKX3.1* promoter, which should be tested for a potential involvement in gene transcription regulation.

**Paper III.** “Epigenetically deregulated genes novel in testicular cancer development include homeobox genes and SCGB3A1 (HIN-1)” Testicular germ cell tumorigenesis may in many ways be looked upon as a genetic and epigenetic caricature of early embryogenesis. In this study, we hypothesized that promoter hypermethylation and subsequent inactivation of genes important in normal early embryogenesis contributed to testicular tumorigenesis. By a candidate gene approach we analyzed the methylation status of a set of homeobox genes *DLX6, EMX2, HOXA9, HOXB5, MSXI*, and *MSX2*, as well as four other selected genes, *CDH13, RASSF1A, RUNX3*, and *SCGB3A1* (alias *HIN-1*) in 7 intratubular germ cell neoplasias and 55 primary TGCTs. In addition, cDNA microarray expression profiles of germ cell tumor cell lines before and after treatment with 5-aza-2’-deoxycytidine were compared. A gene list of 99 potential epigenetic targets was identified from this discovery based global approach. Among these, only the 28 genes with reduced gene expression in a panel of
untreated primary tumors relative to normal testis tissue were considered for further analyses, limiting the number of likely candidates. From the final list \textit{CGGBP1}, \textit{CGRRF1}, \textit{SMARCC2}, \textit{SORBS1}, and \textit{XPA} were analyzed for promoter hypermethylation. In summary, the three most frequently methylated genes were \textit{SCGB3A}, methylated in 54\% of the nonseminomas, \textit{RASSF1A} (29\%), and \textit{HOXA9} (26\%). For a subset of the samples the promoters of \textit{SCGB3A1} and \textit{HOXA9} were cloned and bisulphite sequenced, and the results confirmed the methylation status of \textit{HOXA9}, but revealed some false positives among the \textit{SCGB3A1} MSP positive cases. \textit{CDH13} and \textit{HOXB5} demonstrated methylation at low frequencies, and \textit{EMX2}, \textit{MSX1}, \textit{RUNX3}, and \textit{SORBS1} only rarely. Overall, the nonseminomas were by far more often methylated than were seminomas, and we could also see significant differences between the various nonseminoma subtypes. This study identifies for the first time homeobox genes as epigenetically regulated targets in testicular cancer. Furthermore, we demonstrate that \textit{SCGB3A1} is a novel gene in testicular tumorigenesis, and one of the most frequently methylated genes reported to date in this cancer type.

\textbf{Paper IV.} “A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines” Tumor cell lines are commonly used as experimental tools in cancer research, but their relevance for the \textit{in vivo} situation is debated. In this study we analyzed the promoter hypermethylation status of six tumor suppressor genes in colon cancer cell lines (n = 20) and primary colon carcinomas (n = 53) stratified by MSI and with known ploidy stem line as well as \textit{APC}, \textit{KRAS}, and \textit{TP53} mutation status. Among the cell lines 15\%, 50\%, 75\%, 65\%, 20\% and 15\% showed promoter methylation for \textit{MLH1}, \textit{MGMT}, \textit{p16\textsuperscript{INK4a}}, \textit{p14\textsuperscript{ARF}}, \textit{APC}, and \textit{E-Cadherin}, respectively, whereas 21\%, 40\%, 32\%, 38\%, 32\%, and 40\% of the primary tumors were methylated for the same genes. Both \textit{MLH1} and \textit{p14\textsuperscript{ARF}} showed higher methylation frequencies in MSI than in MSS primary tumors. Moreover, \textit{p14\textsuperscript{ARF}}, which
indirectly inactivates TP53, was methylated more frequently in tumors with normal TP53 than in mutated samples, but the difference was not statistically significant. In addition to determining the methylation frequency of six essential tumor suppressor genes, this paper shows that colon cancer cell lines are in general relevant in vitro models for the in vivo situation, as the cell lines display many of the same molecular alterations as do the primary carcinomas.

**Paper V.** “Identification of Novel Gene Targets in Colorectal Tumorigenesis: ADAMTS1, CRABP1, and NR3C1” In order to identify novel gene targets for inactivation by DNA hypermethylation in colorectal cancers, we compared the gene expression profiles of four colon cancer cell lines before and after treatment with the demethylating agent 5-aza-2’-deoxycytidine using microarrays. Ninety-three array elements responded to treatment and 88 of these were present on oligo microarrays used to analyze the gene expression of a panel of 18 primary colorectal carcinomas and three normal colon tissue samples. The median of 60 of these genes was down-regulated across the panel of tumor samples relative to normal colon tissue and 21 of them were found to contain a CpG island in their 5’ gene region. From this list, we selected four candidates encoding proteins with potential roles in tumor development: ADAMTS1, CRABP1, NDRG1, and NR3C1. The promoters of ADAMTS1, CRABP1, and NR3C1 were hypermethylated in 17/20 (85%), 18/20 (90%), and 7/20 (35%) colon cancer cell lines, respectively, whereas NDRG1 was unmethylated in all samples from the MSP analyses. Bisulphite sequencing confirmed these results and in vitro quantitative gene expression analyses demonstrated that the transcripts of hypermethylated genes were either absent or reduced, statistically significant for two of the three genes. The methylation status of ADAMTS1, CRABP1, and NR3C1 were also investigated in primary colorectal adenomas and carcinomas and 23/63 (37%), 7/60 (12%), and 2/63 (3%) adenomas, and 37/52 (71%), 25/51
(49%), and 13/51 (25%) carcinomas were hypermethylated for the respective genes. In conclusion, this study shows that ADAMTS1, CRABP1, and NR3C1 are novel epigenetically inactivated genes of importance in colorectal tumorigenesis.

**Paper VI.** “DNA methylation in benign and malignant large bowel tumors and their *in vitro* models”. Although CpG island hypermethylation has been demonstrated to be a frequent event during colorectal tumorigenesis, we have limited knowledge regarding the epigenetic alterations influencing the early stages of this process. In paper VI, we therefore analyzed the promoter methylation status of 11 selected genes, including ADAMTS1, APC, CRABP1, HOXA9, MGMT, MLH1, NR3C1, p16INK4a, PTEN, RUNX3, and SCGB3A1, in hyperplastic polyps (n = 12), adenomas (n = 63), colorectal carcinomas (n = 53) and colon cancer cell lines (n = 20). Some of the genes have previously been analyzed by us in parts of the series (paper IV and V) and the results were included here in order to compare the methylation profiles of the benign lesions with that of malignant ones and *in vitro* models. In total, 83% hyperplastic polyps, 73% adenomas, 89% carcinomas, and 100% cell lines were methylated in one or more of the 11 genes analyzed with an average of 4.1, 1.7, 3.4, and 5.3 methylated genes in the four respective tumor groups. The most frequently methylated genes in the primary carcinomas were ADAMTS1, CRABP1, and MGMT. Methylation frequencies among the individual genes typically increased with malignancy, with the apparent exception of hyperplastic polyps. These lesions displayed methylation frequencies comparable to those of carcinomas and may belong to the subgroup of serrated adenomas. In addition, HOXA9, MGMT, and APC showed similar methylation frequencies in adenomas and carcinomas, suggesting that the inactivation of these genes occur early in colorectal tumorigenesis. Further, we confirmed that methylation is most common in carcinomas with MSI and proximal location. Additional associations to gender, age, and polyp size were also identified for some of the methylated genes. In general,
methylated frequencies were higher in cell lines than in primary tumors and statistically significant for \textit{CRABP1}, \textit{p16}^{\text{INK4a}}, and \textit{SCGB3A1}. However, as the overall methylation profiles of the two groups were comparable, colon cancer cell lines can be considered representative epigenetic models for large bowel carcinomas.
Methodological considerations

Modification of DNA - Sodium bisulphite treatment

The sodium bisulphite reaction was initially described in 1970 [326], but was not used for 5-methylcytosine detection until Frommer and Clark described the protocol in detail in the 1990ties [327,328]. This was a methodological revolution for epigenetic research and became the backbone of several reliable techniques used to distinguish between cytosine and 5-methylcytosine in complex genomes. The DNA modification is based on the different sensitivity of cytosine and 5-methylcytosine to deamination by bisulphite under acidic conditions. In this process, unmethylated cytosine undergoes conversion to uracil whereas the unreactive 5-methylcytosine remains a cytosine. Although the efficiency of chemical conversion of unmethylated cytosine to uracil can be close to 100% in an optimal reaction, a 95 - 98% success rate is more common [329]. A less efficient rate of conversion, causing unmethylated cytosines to reside in the bisulphite modified DNA, can create problems when interpreting the results of subsequent DNA methylation analyses (see below). In order to ensure complete DNA modification, the DNA should be of high quality, and fully denatured prior to the bisulphite reaction, since the modification is highly single strand specific. Correct pH and incubation temperature are also crucial for optimal chemical reaction during several steps in the modification process. Furthermore, and to minimize the oxidative degradation (by the automatic oxidization of bisulphite), a free radical scavenger like hydroquinone should be included in the reaction [330]. Various post bisulphite modification DNA methylation analyses have been described, usually sequence-, PCR-, and/or restriction enzyme-based
Discussion

[331], and MSP and bisulphite sequencing, the two methods used in this thesis, are discussed below.

**Methylation specific-PCR**

MSP was first described in 1996 [332] and is now the most widely used technique for studying DNA methylation. After bisulphite treatment of the DNA, PCR is performed using one primer set that will anneal to and amplify methylated fragments only and a second set that will anneal to and amplify unmethylated sequences. The specificity of this assay therefore relies on the match or mismatch of the primer sequence to bisulphite treated DNA and thus, the choice of primers used for MSP can greatly influence the results obtained. A good rule of thumb is to include as many CpG sites as possible in the primer sequences to ensure easy discrimination between methylated and unmethylated sequences during the reaction. A CpG site in the 3`-end of the sense primer further increases the specificity. Finally, primers should include non-CpG cytosines to avoid amplification of false positives. If such cytosines are not included, the primer set specific for methylated DNA might amplify the unmethylated unmodified DNA. Indeed, Rand and co-workers demonstrate that such co-amplification results in an overestimation of DNA methylation [333]. The majority of the MSP primers included in our assays fulfil the criteria mentioned above. The specificity of the few primers that include only one CpG site (like the sense primer of CRABP1) has been tested. These primers are still able to discriminate between unmethylated and methylated alleles and none of them are able to amplify unmodified DNA.

For some genes it has been demonstrated that hypermethylation of specific CpG sites are relevant for gene expression and others are not. One example is the MLH1 gene, in which hypermethylation of eight CpG sites approximately 200 base pairs upstream of the
transcription start point invariably correlates with the lack of gene expression [334,335]. If the MSP primers had been designed to amplify a fragment 400 base pairs further upstream, as good as all samples would have been methylated regardless of gene expression status [334]. When using MSP, one should therefore ensure that the primers anneal with an “expression relevant” part of the promoter. All MSP primers included in paper IV meet this demand, including the MLH1 primers designed by us. However, only a minority of the CpG island containing gene promoters have been mapped in detail in this manner. The next best thing for MSP analyses of such genes, would be to design the primers in close proximity to transcription start (which we have done for the majority of the remaining primers), as the methylation status of CpG sites in this area is more likely to affect gene expression than sites far upstream in the promoter. Analyses of expression should be included in these studies to elucidate whether promoter hypermethylation is associated with lack of gene expression or not (see “validation of the effect of DNA hypermethylation” below).

**Interpretation of MSP results**

The sample products from MSP, separated by gel electrophoresis and stained with ethidium bromide, often show variable band intensities. This is particularly noticeable from the reaction that amplifies methylated alleles when using human tissue samples (e.g. tumor biopsies) as template. Even though the MSP is a non-quantitative method [336], the band intensities are highly reproducible among independent analyses of individual samples. Since MSP is the most sensitive DNA methylation technique available, and can detect as little as one methylated allele in a pool of 1000 unmethylated alleles [332], the intensity variations among samples are after all most likely reflecting different amounts of methylated alleles in the template analyzed. To my knowledge, no guidelines for scoring and interpreting MSP results with various band intensities have been published. For the papers included in this thesis, we
have scored the methylated MSP samples relative to the intensity of the positive control [239]. *In vitro* treated placenta DNA is the positive control for the methylated MSP and is 100% methylated in all CpG sites. Samples with gel band intensities equal or stronger than this positive control are scored as heavily methylated (++), whereas samples with less intense gel bands than the positive control are scored as weakly methylated (+). Samples with no band from the methylated reaction, displaying a band in the unmethylated reaction are scored as unmethylated. This scoring is described in detail in paper I, which also includes an illustrative gel picture of the three categories of samples (heavily methylated, weakly methylated, and unmethylated).

The TGCTs and colorectal carcinomas analyzed throughout this work are from tumor series containing a high fraction of tumor cells. Even though a few of these cells should display promoter methylation of the gene in question, it is highly unlikely that this will affect the carcinoma phenotype, since the majority of tumor cells will still produce the protein. Hence, only carcinomas and cell lines displaying strong gel bands from amplification with the methylation-specific primer set are acknowledged to be hypermethylated. This conservative way of classifying methylated samples limits the number of false positives, as well as relaxes the sample size requirements, since two times two contingency tables, rather than a three times two tables can be used for statistical analyses. Benign lesions on the other hand, are expected to contain a mixture of cells. The admixture of unmethylated DNA from these cells will dilute the neoplastic epithelial DNA and thereby mask the true methylation status. We therefore acknowledge benign tumors with weak - as well as benign tumors with strong – methylation-specific gel bands to be methylated.
Overall, MSP is a rapid method for studying DNA methylation, and the high sensitivity of the method allows analyses of limited sample material, including micro-dissected sample sets \cite{332}. However, care should be taken to optimize the conditions when performing MSP, and it seems necessary to establish a consensus scoring of such results. Finally, MSP gives only a yes-no answer, from which complete methylation, or complete absence of methylation across the region, is assumed. Therefore MSP should be used in combination with bisulphite sequencing analyses, which can establish the methylation status of individual CpG sites.

Methylation profiles obtained by MSP provide information on the methylation status across many sites in the genome. However, it is important to keep in mind that these profiles cannot be compared directly. The methylation frequencies are dependent on the representativeness of the samples series, exemplified by the distribution of MSI tumors versus MSS tumors or of histological subgroups like nonseminomas versus seminomas. The total number of samples included from various classifications will therefore affect the final methylation frequency.

**Bisulphite sequencing**

Bisulphite sequencing can be considered the gold standard of DNA methylation analyses. This method allows a positive display of individual 5-methylcytosines in the gene promoter after bisulphite modification as unmethylated cytosines appear as thymines, while 5-methylcytosines appear as cytosines in the final sequence \cite{328}. In contrast to MSP primers, the primers for bisulphite sequencing must be designed so that they make no distinction between methylated and unmethylated alleles. These primes should therefore anneal to regions with no CpG sites. Moreover, it is important to optimize the PCR conditions so that both the methylated and unmethylated DNA are amplified with equal efficiency \cite{337}.
Two different approaches can be used for bisulphite sequencing the amplified PCR product. First, the amplified PCR product can be sequenced directly. This will result in an average value for methylation of the sample analyzed (e.g. a primary tumor). The relative peak heights of the cytosine to thymine peaks at each CpG site can be used to semiquantitate the methylation levels [338]. Alternatively, the amplified PCR product can be cloned into plasmid vectors followed by sequencing of the individual clones. This will result in a more detailed methylation profile, as the individual clones might contain different degrees of methylation. The majority of tumor samples analyzed contain some normal cells, and the DNA from these cells is most likely unmethylated. These unmethylated normal alleles will be readily distinguishable from the methylated tumor alleles by the cloning approach. However, in direct sequencing the normal alleles will affect the methylation ratio calculated for each CpG site. Moreover, if methylation levels at any one site are lower than 25%, direct sequencing will not be sensitive enough to be reliable for methylation detection, and the cloning approach is recommended in these cases [330].

The bisulphite sequencing method is quite laborious and time consuming, and a good compromise for accurate DNA methylation analyses is to use the rapid MSP for initial methylation screening and subsequently confirm the findings by bisulphite sequencing a subgroup of the samples. This is the approach we have used in several of our papers. In paper III we sequenced individual clones, as we were primarily interested in confirming the methylation status in TGCTs. However, in paper II and V, we used direct bisulphite sequencing, since most of the sequenced samples were cell lines, which are assumed to be a monoclonal population.
Validation of the effect of DNA hypermethylation

Not all hypermethylated CpG island containing genes are inactivated. CpG islands often span more than 1 kilobase of the gene promoter and the methylation status within this region is sometimes mistakenly assumed to be homogenous. In reality, the methylation status varies among the regions within a CpG island [339]. One example of this is the above-mentioned MLH1 and another is RASSF1A. Methylation mapping of the promoter and first exon of RASSF1A using microarray-technology has shown that hypermethylation of a region covering both the transcription and the translation start site is associated with reduced gene expression [340].

These examples underline the importance of expression validation upon the presentation of novel methylated genes. A commonly used approach is reverse transcription-PCR (RT-PCR), which measures the mRNA level transcribed from the gene of interest. In paper V, we used a real time quantitative variant of RT–PCR to validate the effect of the hypermethylation identified in three genes novel to colorectal tumorigenesis. In spite of the relatively small sample set (n = 20), hypermethylation of the three targets was found associated with reduced gene expression, although proven statistically significant only for two of the three genes.

Protein expression is also frequently used to measure the effect of CpG island promoter hypermethylation. However, more samples are generally required to obtain statistically significant associations using measurements at this level, since various post-transcriptional modifications might also affect the expression pattern. Moreover, it can be necessary to develop new and good antibodies if commercial ones are not available, which is time consuming and rather expensive. In paper I we used a tissue microarray with 510 testicular tissue cores from 279 patients diagnosed with TGCT for immunohistochemical analyses of
the MGMT protein. Twenty of these patients overlapped with the DNA used for MGMT methylation analyses and among these an association was seen between hypermethylation and loss of protein expression, although not statistically significant. The MGMT staining results for all 510 testicular tissue cores can be found in a separate paper [216]. By using tissue microarrays, valuable archival tumor material is better exploited than by sectioning the whole paraffin embedded tissue sample, time and costs of immunohistochemical analyses are considerably reduced, and, if connected to a good database, clinical, pathological, and molecular data can be examined in large series.

**Genome wide gene expression analyses in cell lines cultured with and without demethylating agents**

In two of our studies (paper III and V) we have used microarray gene expression in combination with 5-aza-2’-deoxycytidine treatment of cancer cell lines to identify novel gene targets epigenetically inactivated in cancer. The 5-aza-2’-deoxycytidine is a ring analog of 2-deoxycytidine, which is incorporated into newly synthesized DNA. Once incorporated into the DNA, 5-aza-2’-deoxycytidine can form an irreversible covalent complex with DNMT1 and thereby deplete the cell of active enzyme [341,342]. During cellular replication, the DNMT1 depleted cells undergo passive demethylation in which the newly synthesized DNA strand remains hypomethylated. As successive rounds of DNA replication are necessary to achieve complete demethylation [343,344], it is important to administer 5-aza-2’-deoxycytidine to dividing cells. Further, the treatment must last long enough to demethylate the entire cell population. To remove the methylation in cancer cell lines, 5-aza-2’-deoxycytidine can be mixed directly in the cell medium. However, the compound is cytotoxic and the concentration should be balanced so that methylated genes are reactivated without
killing the cells. Furthermore, 5-aza-2’-deoxycytidine is unstable in aqueous solutions, making frequent changes of newly prepared medium necessary.

Our attempts to use this global approach to identify new target genes were awarded differently in TGCTs and colorectal carcinomas, respectively. In colorectal cancer, three of the four identified target genes were hypermethylated, whereas all five candidates in TGCTs were unmethylated. There are several reasonable explanations for this discrepancy in experimental outcome. First of all, in the study of testicular cancer only two germ cell tumor cell lines were cultured in parallel with and without drug treatment, in contrast to the colorectal cancer study in which four cell lines were used. Furthermore, both of the germ cell tumor cell lines originated from embryonal carcinomas, which is the histological subtype of primary tumors displaying the least methylation. Overall, we believe that the strict criteria used for the global approach in the colorectal cancer study might have enhanced the likelihood of detecting true methylated targets. The different rates of success in the two cancer types can also be explained by biological differences, which are dealt with below (see “Disparities and similarities in epigenetic changes in TGCTs and colorectal carcinomas”).

In general, the treatment of cell lines with 5-aza-2’-deoxycytidine results in the activation of numerous genes. In addition to the expected CpG island containing genes, the gene panels also include several targets with unmethylated CpG islands as well as genes with no CpG island in their promoter. APAF-1 is an example of the latter group, in which the RNA and protein levels increased 8- to 20-fold upon 5-aza-2’-deoxycytidine treatment of APAF1-negative melanoma cells. The unmethylated and non-CpG containing genes could be up-regulated as a cellular response to the toxic effect of the 5-aza-2’-deoxycytidine treatment, or by hypermethylated enhancers located elsewhere in the sequence. However, the
exact mechanism by which the drug activates these genes remains unclear. From the colon cancer cell line study, *NDRG1* turned out to have an unmethylated CpG island, whereas all five testicular cancer candidates belonged to the same group.

Several studies analyzing the re-expression of hypermethylated genes are using a nucleoside analog (like 5-aza-2’-deoxycytidine) in combination with a histone deacetylase inhibitor (*e.g.* trichostatin A) [295,349,350]. Histone deacetylase activity is important in the transcriptional repression of methylated sequences [351] and a combination of these two drugs have been shown to cause synergistic reactivation of silenced genes in cancer cell lines [352]. Given the crosstalk between these epigenetic pathways, a combination of these drugs could be more effective than individual drugs alone.

**DNA methylation in cancer development**

It should be mentioned that hypomethylation of genes normally silenced by DNA methylation has been shown to have an oncogenic effect in cancer [110-115]. However, this has been studied to a less extent than the inactivation of tumor suppressor genes or repair genes by hypermethylation. We have focused on the latter, and will discuss this in the context of the two disease models we have studied.

**DNA methylation changes in testicular germ cell tumorigenesis**

Although the characteristic cytogenetic marker, i(12p), has been known for more than two decades, and many genetic changes have been described in TGCTs, rather few gene targets can be considered to be major contributors to the tumorigenesis [193]. At the start of this thesis, the epigenetic knowledge of these tumors was very limited. In our first study we
selected MGMT for DNA methylation analyses because it maps to chromosome band 10q26, which is homologous to the Tgct-1 locus in a mouse model used to study susceptibility of TGCT [353]. Furthermore, MGMT was a known target gene for inactivation by hypermethylation in several other cancer types [132,354]. We showed for the first time that MGMT is hypermethylated in a substantial fraction of TGCTs. Furthermore, hypermethylation of MGMT was strongly associated with nonseminomatous TGCTs, and was thereby the first gene identified by a candidate gene approach to confirm the global methylation differences between seminomas and nonseminomas reported by Smiraglia and co-workers [221]. Later that same year, Koul and co-workers published a list of tumor suppressor genes displaying higher methylation frequencies in nonseminomas than in seminomas [223]. They confirmed MGMT as a target gene inactivated by hypermethylation in TGCTs and also identified new epigenetic targets, like RASSF1A and HIC1. Moreover, they demonstrated that the various histological nonseminoma subtypes displayed different methylation frequencies of the gene panel analyzed. The highest frequency was found among yolk sac tumors, whereas teratomas displayed the least methylation [223]. This methylation distribution among histological subtypes is different from our findings (paper III), in which teratomas display more, and the embryonal carcinomas display significantly less methylation than do the rest of the nonseminomas. The inconsistent methylation distribution in the two studies might be related to biological differences in the tumor series analyzed, gene specific methylation differences among histological subtypes, or tissue heterogeneity in the individual TGCTs analyzed. The TGCTs often contain both seminomatous and nonseminomatous components or a mixture of various nonseminomatous histologies. Since the presence of normal cells (lymphocyte infiltration) or other histological tumor subtypes might mask the true methylation status, careful examination of tumor histology and subsequent analyses of pure tissue components is preferable. In paper III, frozen tissue sections stained with
hematoxilin and eosin (HE) demonstrated that the majority of samples analyzed fulfilled this criterium [192].

Testicular germ cell tumorigenesis may in many ways be looked upon as a genetic and epigenetic caricature of early embryogenesis [355], which is supported by the histological subtype-specific epigenotypes we identified in paper III. These epigenotypes seem to parallel the epigenetic reprogramming of the early embryonic development. We also found hypermethylation of specific homeobox genes, like \textit{HOXA9} and \textit{HOXB5}. This is the first time epigenetic inactivation of members from this gene family has been reported in TGCTs and might indicate that developmental genes play an important role in the testicular tumorigenesis. However, five other homeobox genes analyzed (including \textit{NKX3.1} from paper II) were unmethylated in the majority of TGCTs, suggesting that specific genes in this family are targeted during the development of abnormal primordial germ cell fate.

\textbf{DNA methylation changes in colorectal tumorigenesis}

Several hypermethylated tumor suppressor genes have been identified in primary colorectal carcinomas (table 3), indicating an important role for epigenetic inactivation in colorectal tumorigenesis. Indeed, except for \textit{PTEN}, all genes analyzed in paper IV and VI were hypermethylated, although at different frequencies. Inactivation of \textit{PTEN} has previously been observed in a variety of sporadic cancers [300] and has been demonstrated to be commonly hypermethylated [356,357]. However, most of the primers designed for DNA methylation studies of this gene anneal to and amplify the \textit{PTEN} pseudogene, which has 98\% sequence homology with \textit{PTEN} [300]. When we used primers specifically designed to amplify the protein encoding \textit{PTEN} gene [358], both colon cancer cell lines and colorectal carcinomas were unmethylated.
In addition to establishing the DNA methylation profile of a large panel of genes with important functions in cancer (summarized in figure 1 in paper VI), their potential associations with clinicopathological data (see below), as well as their relevance to in vitro models were examined. We also identified three novel targets inactivated by promoter hypermethylation in colorectal tumorigenesis, *ADAMTS1*, *CRABP1*, and *NR3C1*, from the global gene expression analyzes of cell lines cultured with and without demethylating agents (paper V). Genetic and epigenetic alterations play a pathogenic role in driving colon neoplasms through the adenoma-carcinoma sequence [250]. Although CpG island hypermethylation has been demonstrated to be an early and frequent event during colorectal tumorigenesis, a limited number of epigenetic changes has been identified in polyps [272,359]. In paper VI, only genes hypermethylated in the primary carcinomas were selected for DNA methylation analyses in colorectal polyps. The individual methylation frequencies varied from gene to gene, but all candidates submitted to analyses were hypermethylated. Hence, this is the first work demonstrating hypermethylation of *ADAMTS1*, *CRABP1*, *HOXA9*, *NR3C1*, *RUNX3*, and *SCGB3A1* in noninvasive lesions, thereby supporting the observation that aberrant hypermethylation of genes occurs early in the tumorigenic process in the large bowel. Inactivation of *HOXA9* in addition to *MGMT* and *APC* might be particularly important for the transformation of adenoma cells, as the methylation frequencies in these lesions are the same as in invasive tumors. Indeed, inactivation of both *APC* and *MGMT* has previously been shown in aberrant crypt foci, which are postulated to be the earliest precursor lesion in colorectal carcinogenesis [360,361]. Among the benign colorectal lesions, hyperplastic polyps displayed significantly higher methylation frequencies than did adenomas. As these non-invasive lesions display distinct epigenetic profiles, they might give rise to subgroups of carcinomas with different molecular changes. Hyperplastic polyps have
Discussion

long been considered benign lesions with no malignant potential, however, recent findings suggest that a subgroup of these lesions, the sessile serrated adenomas, can progress into malignancy [249,362-364]. These tumors are typically larger than the true hyperplastic polyps, arise in the proximal colon, show a high mutation rate in the \textit{BRAF} gene [365], and may give rise to both MSS and MSI carcinomas, depending on the nature of the accumulated molecular changes [249,364].

Regarding overall CpG island hypermethylation, cancer cell lines have in general demonstrated an increased frequency of hypermethylation compared with primary tumors [366]. However, only a limited number of the genes analyzed have shown a statistically significant difference in methylation frequency [366]. Findings in both paper IV and VI indicate that the genetic and epigenetic changes in colorectal carcinomas and colon cancer cell lines are overall comparable. Indeed, among several cancer types examined, colon cancer cell lines have been shown to resemble the most their respective primary tumor regarding DNA methylation [367]. Colon cancer cell lines are therefore relevant \textit{in vitro} models for the \textit{in vivo} situation, in spite of the increase in average number of genes methylated per sample from primary tumors to cell lines.

We identified several clinicopathological variables related to hypermethylation in studies IV and VI. Approximately half of the genes analyzed were more frequently methylated in the MSI than in the MSS carcinomas. Previous studies have reported similar skewed distributions [319-321], and the same association has also been identified in non-malignant precursor lesions [272]. As MSI tumors are strongly associated with proximal tumor location [260,368], it is not unexpected that we see an association between hypermethylation and right-sided tumor location in the colon. These two associations do, however, indicate that DNA
hypermethylation plays a more important role in proximal colon tumorigenesis than in distal colon tumorigenesis. Associations have also been found between hypermethylation and female gender. Furthermore, for ADAMTS1 we found increased methylation frequencies with increased adenoma size, and a similar trend could be seen for HOXA9 and p16\(^{INK4a}\) methylation, although not statistically significant.

**Disparities and similarities in epigenetic changes in TGCTs and colorectal carcinomas**

The molecular genetics of TGCTs and colorectal carcinomas is fundamentally different [369,370]. This is underlined by examples like TP53, which is frequently mutated in a variety of human malignancies, including in colorectal cancer [371], but are rarely mutated in TGCTs [211]. Significant differences between these tumor types are also evident at the epigenetic level. The cell cycle inhibitor p16\(^{INK4a}\) is frequently hypermethylated in colorectal carcinomas [322] as well as in a wide variety of other human primary tumors and cell lines [127,128]. Even though p16\(^{INK4a}\) rivals the TP53 gene in inactivation frequency in many tumor types, it is unmethylated in all TGCTs [239].

From our DNA methylation analyses, TGCTs and colorectal carcinomas are shown to have quite different methylation profiles. Whereas hypermethylated genes are frequently found among tumors of the latter group, they are more rare within the first group. From table 2 and 3 we see that approximately 35 genes are hypermethylated at medium or high frequencies in colorectal carcinomas (when genes inactivated by age-specific methylation is excluded), whereas only a fifth are hypermethylated at similar extents in TGCTs. Some of this difference might be due to the uneven research attention these two cancer types receive. Whereas colorectal cancer is intensively studied worldwide, fewer research groups are working with
testicular cancer. However, different tumor biology is more likely to be the main contributor to the methylation profile differences. Ninety-two percent (12/13) of the genes we analyzed for DNA methylation in colorectal carcinomas were methylated, whereas only 44% (8/18; when disregarding sample 2110) were methylated in TGCTs. This difference in hypermethylation prevalence might also explain our lack of success when we tried to identifying new epigenetically inactivated target genes in TGCTs. Findings by Smiraglia and colleagues support this, as TGCTs have relatively low hypermethylation frequencies while colorectal carcinomas have relatively high frequencies from RLGS studies [372]. This can also be seen in the precursor lesions. In contrast to the adenomas of the colorectum, the non-invasive precursor of TGCTs, the intratubular germ cell neoplasias, are only rarely hypermethylated.

Despite the overall epigenetic contrasts seen between TGCTs and colorectal carcinomas, similarities can be found at the single gene level. MGMT, HOXA9, and SCGB3A1 are hypermethylated in the development of both cancer models – arising in primordial germ cells and epithelial cells, respectively, and these genes are therefore more likely to be of general importance to cancer than are targets restricted to the individual diseases. Indeed, MGMT is a DNA repair protein that protects the cells against mutagenesis and malignant transformation [373]. The function of SCGB3A1 remains to be established, whereas HOXA9 is a member of the homeobox gene family, encoding proteins that regulate morphogenesis and cell differentiation during embryogenesis [374].

**Clinical applications of epigenetic changes in cancer**

The main aim of this thesis has been to learn more about DNA methylation in cancer development. This mechanism, along with other epigenetic changes in cancer, can be of great
clinical value. In contrast to the genetic changes, which are irreversible, epigenetic changes can be reversed, and are therefore interesting therapeutic targets. Although the clinical applications of DNA methylation have not been the focus of our studies, I will briefly refer the current state of the art and discuss some of the clinical implications in the rapidly advancing field of epigenetics.

**Gene promoter hypermethylation as a biomarker for early detection, intervention, and prognosis**

Over the last decade, molecular markers have emerged as promising tools for early cancer detection, patient management, and assessment of prognosis. In addition to being quite frequent, DNA methylation changes have been reported to occur early in tumorigenesis and are therefore potentially good early indicators of existing disease [290,375]. In addition to bodily fluids, DNA methylation can be analyzed from detached tumor cells present in *e.g.* stool, urine, or sputum samples. At the present, several reports have examined various hypermethylated biomarkers for non-invasive tests for early cancer detection, and some of them are quite promising (reviewed in [7] and [376]). Hypermethylation of *DAPK1*, *MLH1*, and *p16INK4a* has been found in serum of 21%, 33%, and 30% of colorectal cancer patients, respectively [377-379]. Additionally, 77% of faeces samples from colorectal cancer patients contained hypermethylated *SFRP2* [380]. As the efficacy of a biomarker assay is determined by its sensitivity and specificity, multiple biomarkers should be included in each assay. Such panels will additionally provide more diagnostic information than single marker assays, like the specific tumor type present in a patient based on a positive blood sample.

Epigenetic changes of genes like *p16INK4a*, *MGMT*, *DAPK1*, and *RASSF1A* have been extensively studied in sputum from cancer-free individuals who were at risk of developing
lung cancer because of smoking [381]. Several of the individuals harboring these changes developed lung cancer after approximately one year of follow up [382,383]. The relative risk for lung cancer is likely to increase along with the increased number of hypermethylated genes detected in sputum from these individuals. Belinsky and colleagues are presently using a case-control study to assess the threshold level for intervention for this group of high-risk individuals [381]. Once individuals pass the intervention point, spiral computed tomography and/or bronchoscopy would be recommended to identify early lung cancer.

The DNA methylation status of individual CpG containing gene promoters can also be used to assess general prognosis. Associations between hypermethylation of specific genes and patient prognosis have been reported for various cancer types [384-389]. DNA methylation has also been reported to predict response to a particular therapy. Glioma patients with hypermethylation of MGMT have increased survival upon treatment with alkylating agents compared with patients with no methylation [354,390,391]. Moreover, by repeated in vitro drug exposure, melanoma cells can acquire resistance to the anti-neoplastic alkylating compound fotemustine by reactivation of the MGMT gene [392].

**Epigenetic silencing mediated by CpG island hypermethylation as a potential therapeutic target**

Epigenetic therapy involving the use of DNA methylation inhibitors and histone deacetylase inhibitors is currently receiving much attention [393]. These agents have been proposed to have anti-tumor properties as they can reactivate the expression of epigenetically silenced tumor suppressor genes [394]. By reactivating key genes in essential molecular signaling pathways, normal cellular functions might be reinitiated in malignant cells [395]. Reactivation of p16INK4a and RB1 can inhibit uncontrolled cell growth [125,128], active DAPK1 and


CASP8 can induce apoptosis [396,397], reactivation of CDH1, SERPINB5, and TIMP3 might prevent metastasis by increasing cell adhesion [398-400], and reactivation of MYOD1 can contribute to cellular differentiation [401]. Moreover, loss of mismatch repair function due to hypermethylation of MLH1 can result in resistance in vitro to a number of clinically important anticancer drugs [402]. 5-aza-2'-deoxycytidine treatment can restore sensitivity to these chemotherapeutic compounds [402] (Figure 9).

![Figure 9. Gene targets of epigenetic therapy.](image)

Demethylating agents, like 5-aza-2'-deoxycytidine (clinically referred to as decitabine), have been extensively used for in vitro studies (e.g. this thesis) and have recently undergone several clinical trials [403,404]. These drugs seem to be particularly efficient in hematological malignancies [403-405] and in chronic myelogenous leukemia [406]. Histone deacetylase inhibitors have also shown potential for the treatment of hematological malignancies in vitro [407] as well as in vivo [408]. By combining demethylating agents with histone deacetylase inhibitors synergistically effects can be achieved [352,409].
Despite the encouraging responses in hematological malignancies, demethylating treatment of solid tumors has given uniformly disappointing results [410-415] including in nonseminomatous testicular cancer [416] and colorectal cancer [417]. Overall, solid tumors display frequent gene inactivation by CpG island DNA hypermethylation, so the lack of 5-aza-2’-deoxycytidine effect is not caused by lack of aberrant methylation [130]. It remains to be seen if epigenetic compounds combined with other drugs will have a better effect in solid tumors.

Epigenetic therapy seems to hold great promises for cancer treatment. However, there are several concerns regarding the clinical application of these agents. First of all, the reactivation of hypermethylated sequences is nonspecific. This means that the therapy also targets the transposable elements, tissue specific –, and imprinted –genes, which are usually inactivated in the normal cell [393]. Furthermore, even though decreased DNA methylation can protect against specific cancer types, like intestinal tumors, it can simultaneously lead to an increased cancer risk in other tissues [102].
Conclusions

We have successfully established the necessary assays for analyzing DNA methylation in our laboratory, including *in silico* analyses to identify CpG islands and design primers, bisulphite conversion of DNA, methylation specific-PCR, bisulphite sequencing, and drug treatment of cell lines to remove CpG methylation genome-wide. The combined candidate gene approach and genome studies have been used to gain insights into the epigenetic development of testicular- and colorectal- cancer. Even though the discovery based microarray approach was more successful in the colorectal cancer study than in the testicular cancer study, new gene targets were identified for both disease models using a candidate gene approach. In conclusion, we show that cancer can be described as an epigenetic- as much as a genetic- disease.

The cytogenetic marker for TGCTs, isochromosome 12p, has been known since the early 80-ties and multiple chromosomal aberrations have been described in these tumors that are typically in the triploid range. The majority of potential target genes for these chromosomal losses and gains still remain unknown. Many candidate genes have been submitted to mutation analyses and although some mutations have been identified, the overall frequencies are quite low. In contrast, we and others have shown that the epigenetic changes are present in rather high frequencies. The results, on which this thesis is based, have identified *MGMT*, *SCGB3A1*, and *HOXA9* as novel targets in testicular germ cell tumorigenesis inactivated in a substantial amount of these tumors. In addition, we have confirmed the involvement of *RASSF1A* and rejected the hypothesis of inactivation of p16<sup>INK6a</sup> through methylation in these tumors. Furthermore, we have identified gene targets such as *HOXB5* and *CDH13* potentially contributing to the progression of a small percentage of these tumors. Seminomas and
Conclusions

nonseminomas seem to have different tumor biology, and we confirmed that nonseminomas display high DNA methylation frequencies of selected target genes whereas seminomas in general are devoid of methylation. Moreover, the gene methylation frequencies vary along the embryonal and extra-embryonal differentiation lineages of germ cell tumors, indicating that methylation changes in testicular tumorigenesis mimic the epigenetic reprogramming during early embryogenesis.

Epigenetic changes are among the earliest changes in colorectal cancer development and can be identified even in very small polyps. We have demonstrated several novel epigenetic targets in benign lesions from the large bowel, including HOXA9, RUNX3, SCGB3A1, ADAMTS1, CRABP1, and NR3C1. We initially identified the latter three as commonly methylated in primary colorectal carcinomas. The malignant tumors display higher methylation frequencies than do adenomas, with the apparent exception of HOXA9 as well as APC and MGMT. Because these genes have comparable methylation frequencies in adenomas and carcinomas, we speculate that their inactivation might be particularly important for the transformation into malignancy. Indeed, previous reports have pinpointed APC and MGMT as early changes in tumorigenesis. Hence, APC, HOXA9 and MGMT may be suitable as biomarkers for early detection of colorectal cancers. For several of the target genes, DNA methylation status is associated with clinicopathological features, including MSI-status, localization, sex and degree of differentiation, which may aid in the future molecular assisted sub-classification of the existing clinical staging. Finally, colon cancer cell lines are relevant in vitro models for the in vivo situation.

It is important to keep in mind that even though the novel target genes presented here are methylated at high frequencies, it remains to be established whether these changes are
causative or merely a consequence of tumorigenesis. Identification of methylated target genes in the development of both cancer models, like $MGMT$, are more likely to be of general importance to cancer than those restricted to the individual diseases.
The importance of epigenetics in tumor development and its potential for clinical implementation, both in terms of therapy and diagnostics, are the basis for the planned continuance of our studies. Below is a brief description of some of the short- and long-term plans and thoughts.

**Expanding our methodological repertoire for epigenetic studies**

Conventional methods used to determine methylation status, like MSP and bisulphite sequencing, are excellent to analyze individual genes, but are unable to evaluate methylation changes on a genome-wide scale. In addition to the microarray approach we describe in paper III and V, several methods have been described for large-scale DNA methylation analyses. Some of them are array based [418-423] whereas others are not [424,425]. The opportunity to analyze the methylation status of several genes simultaneously is tempting, and we are considering establishing such a method in our lab. However, at the present there are strong pros and cons with each of these methods [339], which will be reviewed in detail before a final choice is made.

**Studies at the chromatin level**

Acetylation and methylation of residues in the tails of nucleosomal core histones has a crucial role in chromatin packing and gene expression [426]. These modifications can be analyzed by chromatin immunoprecipitation (ChIP), a method we plan to implement and to combine with DNA microarray technology (ChIP-chip). This will provide a powerful high-throughput
method of genome-wide mapping of protein – DNA interactions in vivo, which can be used to identify novel epigenetically regulated genes in cancer. With ChIP, we can also analyze whether or not known DNA methylation changes are reflected in the modification of histone.

Furthermore, we will address the chromatin organization in a large cohort of colorectal cancer patients by texture analyses of formalin-fixed, paraffin embedded material. Advanced computerized image informatics has been developed for these analyses, which will be performed in another lab.

**Functional studies**

Functional assays in combination with methylation analyses can shed light on whether the individual DNA methylation changes are causes or consequences of tumorigenesis. Even though numerous epigenetic changes already have been identified, we’re still probably only looking at the tip of the iceberg. We plan to establish some functional assays in our own lab, but we will also rely on cell biology expertise from collaborative groups.

**Project plans for epigenetic studies in our disease models**

**Epigenetics in premalignant lesions and normal tissue**

The possibility of easily identifying and surgically removing early non-invasive- as well as more advanced- lesions makes colorectal cancer a unique disease model for tumor developmental studies. We will continue to analyze DNA methylation changes in colorectal polyps in order to design a panel of biomarkers that can be used for a noninvasive test for early detection. We will use the same approach as described in paper V and treat colon cancer cell lines with demethylating – and histone deacetylating - agents, and thereafter analyze the
effect on the gene expression. The cell lines will probably contain more changes than do the polyps, and in order to increase the success rate, we will compare the cell line findings with global gene expression analyses of polyps and adjacent normal colon tissue.

Furthermore, we have planned a detailed methylation profiling of normal tissue samples. Normal mucosa, taken from the same large bowels that we have cancer material from, will be matched by age and sex and compared with mucosa from individuals without colorectal cancer (samples removed due to other diseases, or post mortem samples). These results will be interesting per se, but also highlight whether DNA methylation findings among benign and malignant samples are tumor specific or not.

**Epigenetic master keys in cancer – do they exist?**

In order to identify potential targets epigenetically deregulated across several cancer diseases, we have established an international collaborative study in which we will use the same global approach as outlined in the present thesis to analyze multiple in vitro models and corresponding primary tumors. By performing these studies in one lab, we can compare the results directly, and identify gene lists of commonly methylated cancer genes (according to certain chosen cut-offs). Here we define “master key” as a gene that is altered in a certain number of the investigated cancer types. Already, we know that the cyclin-dependent kinase inhibitor $p16^{INK4a}$ is inactivated by hypermethylation in more than half of the tumor types analyzed [132].
Epigenetically regulated genes in sporadic cancers of the types found in the HNPCC tumor spectrum

Many mechanisms and gene targets are common among the sporadic equivalents to HNPCC tumors. In order to compare the epigenetic changes in these tumors, we have an ongoing study in which we analyze the DNA promoter methylation of individual genes in colorectal-, gastric-, endometrial-, ovarian-, and pancreatic- cancer. The tumors from each disease are stratified according to MSI status. This subproject is partly overlapping with the study mentioned in the previous paragraph.

Validation analyses

Tissue microarrays are excellent tools for clinical validation of biomarkers. We have already constructed such arrays for all our disease models (colorectal carcinomas and polyps, TGCTs, and malignant peripheral nerve sheath tumors; MPNST; not included in the present thesis) [216,427](and unpublished data). By using tissue microarrays, we can validate the impact of specific targets altered by epigenetic or other mechanisms, in hundreds of samples simultaneously, in a cost and time saving manner. We have already shown the usefulness of this approach in the TGCT model [216]. The quality of the output is highly dependent on that of the input, and all tumor samples should be thoroughly examined by a pathologist prior to tissue microarray construction. After all, the individual tumor core biopsies included in the array are very small. Finally, in silico analyses, including cross-examination of private and publically available data will be performed to design experimental studies for functional validation.
# REFERENCES


13. MR Teixeira, S Heim: Multiple numerical chromosome aberrations in cancer: what are their causes and what are their consequences? *Semin Cancer Biol* 2005, **15**: 3-12.


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


43. VG Allfrey: Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. *Proc Can Cancer Conf* 1966, **6**: 313-335.


59. TM DeChiara, EJ Robertson, A Efstratiadis: **Parental imprinting of the mouse insulin-like growth factor II gene.** *Cell* 1991, **64**: 849-859.


64. AP Bird, AP Wolffe: **Methylation-induced repression--belts, braces, and chromatin.** *Cell* 1999, **99**: 451-454.


69. SK Murphy, RL Jirtle: **Imprinting evolution and the price of silence.** *Bioessays* 2003, **25**: 577-588.

70. DP Barlow: **Methylation and imprinting: from host defense to gene regulation?** *Science* 1993, **260**: 309-310.

71. JH Thomas: **Genomic imprinting proposed as a surveillance mechanism for chromosome loss.** *Proc Natl Acad Sci U S A* 1995, **92**: 480-482.

72. S Varmuza, M Mann: **Genomic imprinting--defusing the ovarian time bomb.** *Trends Genet* 1994, **10**: 118-123.


76. CJ Brown, A Ballabio, JL Rupt, RG Lafreniere, M Grompe, R Tonlorenzi, HF Willard: **A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome.** *Nature (Lond)* 1991, **349**: 62-64.

References


97. JA Yoder, CP Walsh, TH Bestor: Cytosine methylation and the ecology of intragenomic parasites. 

98. T Takahara, T Ohsumi, J Kuromitsu, K Shibata, N Sasaki, Y Okazaki, H Shibata, S Sato, A Yoshiki, M 
Kusakabe et al.: Dysfunction of the Orleans reeler gene arising from exon skipping due to 
transposition of a full-length copy of an active L1 sequence into the skipped exon. _Hum Mol Genet_ 

99. AE Peaston, AV Eviskov, JH Graber, WN de Vries, AE Holbrook, D Solter, BB Knowles: 
Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. _Dev Cell_ 

100. A Buzdin, E Gogvadze, E Kovalskaya, P Volchkov, S Ustyugova, A Illarionova, A Fushan, T 
Vinogradova, E Sverdl: The human genome contains many types of chimeric retrogenes generated 


102. F Gaudet, JG Hodgson, A Eden, L Jackson-Grusby, J Dausman, JW Gray, H Leonhardt, R Jaenisch: 

103. RZ Chen, U Pettersson, C Beard, L Jackson-Grusby, R Jaenisch: DNA hypomethylation leads to 

104. H Tsuda, T Takarabe, Y Kanai, T Fukutomi, S Hirohashi: Correlation of DNA hypomethylation at 
pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and 

105. WA Schulz, JP Elo, AR Florl, S Pennanen, S Santourlidis, R Engers, M Buchardt, HH Seifert, T 
Visakorpi: Genomewide DNA hypomethylation is associated with alterations on chromosome 8 in 

106. C Lengauer, KW Kinzler, B Vogelstein: DNA methylation and genetic instability in colorectal cancer 

107. M Ehrlich: The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency 

108. GZ Qu, PE Grundy, A Narayan, M Ehrlich: Frequent hypomethylation in Wilms tumors of 

109. MF Chan, R van Amerongen, T Nijjar, E Cuppen, PA Jones, PW Laird: Reduced rates of gene loss, gene 
silencing, and gene mutation in Dnmt1-deficient embryonic stem cells. _Mol Cell Biol_ 2001, 21: 7587-
7600.

110. AP Feinberg, B Vogelstein: Hypomethylation of ras oncogenes in primary human cancers. _Biochem 

111. M Hanada, D Delia, A Aiello, E Stadtmauer, JC Reed: bcl-2 gene hypomethylation and high-level 

112. M Nakayama, M Wada, T Harada, J Nagayama, H Kusaba, K Ohshima, M Kozuru, H Komatsu, R Ueda, 
M Kuwano: Hypomethylation status of CpG sites at the promoter region and overexpression of the 

113. N Nakamura, K Takenaga: Hypomethylation of the metastasis-associated S100A4 gene correlates 
with gene activation in human colon adenocarcinoma cell lines. _Clin Exp Metastasis_ 1998, 16: 471-
479.


135. KD Robertson, PA Jones: The human ARF cell cycle regulatory gene promoter is a CpG island which can be silenced by DNA methylation and down-regulated by wild-type p53. Mol Cell Biol 1998, 18: 6457-6473.


References


References


204. AM Ottesen, NE Skakkebaek, C Lundsteen, H Leffers, J Larsen, E Rajpert-De Meyts: High-resolution comparative genomic hybridization detects extra chromosome arm 12p material in most cases of carcinoma in situ adjacent to overt germ cell tumors, but not before the invasive tumor development. Genes Chromosomes Cancer 2003, 38: 117-125.


249. CV Wynter, MD Walsh, T Higuchi, BA Leggett, J Young, JR Jass: Methylation patterns define two types of hyperplastic polyp associated with colorectal cancer. *Gut* 2004, **53**: 573-580.


265. AR Florl, C Steinhoff, M Muller, HH Seifert, C Hader, R Engers, R Ackermann, WA Schulz: Coordinate hypermethylation at specific genes in prostate carcinoma precedes LINE-1 hypomethylation. *Br J Cancer* 2004, **91**: 985-994.


269. JM Cunningham, ER Christensen, DJ Tester, CY Kim, PC Roche, LJ Burgart, SN Thibodeau: Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res* 1998, 58: 3455-3460.


283. T Nagasaka, H Sasamoto, K Notohara, HM Cullings, M Takeda, K Kimura, T Kambara, DG MacPhee, J Young, BA Leggett et al.: Colorectal cancer with mutation in BRAF, KRAS, and wild-type with


References


References


Paper I

B Smith-Sørensen, GE Lind, RI Skotheim, SD Fosså, Ø Fodstad, A-E Stenwig, KS Jakobsen, and RA Lothe

Frequent promoter hypermethylation of the $O^6$-Methylguanine-DNA Methyltransferase ($MGMT$) gene in testicular cancer

*Oncogene.* 2002 Dec 12;21(57):8878-84
Frequent promoter hypermethylation of the *O*-Methylguanine-DNA Methyltransferase (MGMT) gene in testicular cancer

B Smith-Sorensen¹, GE Lind², RI Skotheim², SD Fosså³, Ø Fodstad¹, A-E Stenwig⁴, KS Jakobsen⁵ and RA Lothe*²

¹Department of Tumour Biology, Institute for Cancer Research, The Norwegian Radium Hospital, 0310 Oslo, Norway; ²Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital, 0310 Oslo, Norway; ³Department of Oncology and Radiotherapy, The Norwegian Radium Hospital, 0310 Oslo, Norway; ⁴Department of Pathology, The Norwegian Radium Hospital, 0310 Oslo, Norway; ⁵Division of Molecular Biology, Biological Institute, University of Oslo, 0310 Oslo, Norway

Testicular germ cell tumours are classified into two major histological subgroups, seminomas and nonseminomas. All tumours display several recurrent chromosomal aberrations, but few target genes have been identified. Previous studies have shown that genome-wide hypermethylation of CpG islands is significantly more prevalent in nonseminomas than in seminomas. We have studied two potential target genes in testicular cancer. A series of 70 tumours were analysed for methylation of CpG sites in the *O*-methylguanine-DNA methyltransferase (MGMT) gene promoter, and in exon 1 of the cyclin-dependent kinase inhibitor 2A gene (CDKN2A). In addition, eight microsatellite markers within and flanking these genes at chromosome arms 10q and 9p, respectively, were analysed for allelic imbalances. Allele alterations were frequently seen at 9p loci (47 out of 70, 67%), but none of the tumours (none out of 55) showed methylation of CDKN2A. On the other hand, a high frequency of MGMT promoter methylation (32 out of 69, 46%) was found, as well as allelic imbalances at 10q markers (50 out of 70, 71%). A significantly higher methylation frequency was found in nonseminomas (24 out of 35, 69%) compared to seminomas (eight out of 33, 24%) (P = 0.0003, Fisher’s exact test). Immunohistochemical analysis of the MGMT protein in a subgroup (n = 20) of the testicular tumours supported the hypothesis of gene silencing being the functional consequence of the promoter methylation. In summary, our data suggest that inactivation of MGMT contributes to development of nonseminomatous testicular cancer.


**Keywords:** CpG methylation; MGMT; CDKN2A; testicular cancer; germ cell tumour; allelic imbalance

---

**Introduction**

Testicular cancer is the most common cancer among young and adolescent Caucasian men and is usually of germ cell origin. Testicular germ cell tumours (TGCT) are divided into two major histological subgroups, seminomas and nonseminomas, both suggested to arise from the precursor stage carcinoma in situ. Nonseminomas contain either embryonal carcinoma, teratoma, yolk sac elements, choriocarcinoma, or a mixture of these components. Today, more than 90% of patients with newly diagnosed TGCT are cured from the disease, and only 10 to 20% of patients with metastatic tumours at time of diagnosis do not become permanently tumour-free (Josefsen et al., 1993).

In addition to the characteristic cytogenetic aberration isochromosome 12p, a number of recurrent DNA sequence changes and chromosome aberrations have been found in TGCT (Atkin and Baker, 1982; Castedo et al., 1989a,b; Heidenreich et al., 2000; Lothe et al., 1993; Murty and Chaganti, 1998; Skotheim et al., 2001a). Nevertheless only few genes contributing to the development of this disease have been identified (Looijenga et al., 1999; Kraggerud et al., 2002; Skotheim et al., 2002).

Methylation imbalance is one of several characteristic features of tumorigenesis. Hypermethylation of CpG sites in promoter regions may result in transcriptional silencing of the genes in question (Baylin and Herman, 2000; Baylin et al., 1998). Genome-wide studies of tumour-associated methylation by restriction landmark genome scanning (RLGS), have revealed a significant difference between seminomas and nonseminomas in the overall methylation pattern (Costello et al., 2000; Smiraglia et al., 1999, 2002). The tumour suppressor gene, CDKN2A (p16⁰INK4A), and the repair enzyme gene, MGMT, are both known targets for methylation inactivation in other cancer types, but their role in TGCT is not known (Esteller, 2000).

The incidence of hypermethylation of the CpG island in the 5’ region of the CDKN2A gene range from 20 to 40% among different types of solid tumours. Together with the incidence of homozygous
deletions, this gene rivals TP53 as the most frequently inactivated tumour suppressor gene in cancer (Baylin et al., 1998). The p16INK4a protein inhibits the catalytic activity of CDK/cyclin D complexes and is thereby involved in the control of cell cycle progression from G1 to S-phase (Serrano et al., 1993). The CDKN2A gene maps to 9p21 (Kamb et al., 1994; Nobori et al., 1994), and loss of heterozygosity at 9p21 loci, implying loss of one CDKN2A allele, has been reported for TGCT (Heidenreich et al., 1998; Murty et al., 1994). Whether 5′-CpG methylation may contribute to silence the CDKN2A gene in TGCT remains unclear (Chau- bert et al., 1997; Heidenreich et al., 1998). To our knowledge methylation of MGMT has not been investigated in TGCT. Cells are normally protected from promutagenic O6-methylguanine alkylation by the action of MGMT (Pegg et al., 1995). MGMT maps to chromosome band 10q26, a region corresponding to the Tgct-1 locus in a mouse model used to study susceptibility of TGCT (Matin et al., 1999).

We have addressed the potential involvement of MGMT and CDKN2A in testicular tumorigenesis by performing 5′-CpG methylation and allelic imbalance (AI) assays in a series of primary tumours.

Results

CpG methylation of MGMT and CDKN2A

The analysed CpG sites upstream of the coding region of MGMT were methylated in 32 out of 69 tumour samples, 46% (Figure 1). The amount of PCR products generated by the methylation-specific primer set varied among the different samples. Samples scored as + and ++ relative to the positive control are illustrated in Figure 1. Tumour specimens scored as methylation positive samples gave sometimes also rise to specific unmethylated products. Both heterogeneity within tumour samples as well as infiltrating normal cells can contribute to this outcome. The number of methylated MGMT samples was significantly higher in nonseminomas (24 out of 35, 69%) compared to seminomas (eight out of 33, 24%). This finding was independent of scoring method (+ or ++), see Table 1. Five of the eight positive seminoma samples with methylated MGMT were taken from pure seminomas, whereas three were from nonseminomas containing seminomatous elements. Only the latter cases showed methylation at a ++ level within the group of seminomas. As expected, no product was obtained from the methylation-specific PCR analysis of the normal testis sample used as negative control. Neither methylation-specific PCR products nor products from unmethylated targets were obtained from the carcinoma in situ sample collected from a patient with seminoma. The other carcinoma in situ sample obtained from a patient with nonseminoma had no methylation of the MGMT gene.

The results of methylation analysis at CpG sites in the CDKN2A gene were scorable for 55 TGCT samples, including 25 samples with methylated sites in MGMT (consisting of 19 nonseminomas and six

---

**Table 1** Incidence of MGMT CpG methylation in TGCT

<table>
<thead>
<tr>
<th>Patients</th>
<th>Methylated Cpg ++</th>
<th>Methylated Cpg +</th>
<th>Unmethylated Cpg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonseminoma</td>
<td>15</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Seminoma</td>
<td>3</td>
<td>5</td>
<td>25</td>
</tr>
</tbody>
</table>

There is a significant difference between number of samples with Cpg methylation among nonseminomas compared to seminomas \(^{a}P=0.002, {b}P=0.0003\), Fisher’s exact test
seminomas). No methylation of CpG sites in CDKN2A was seen in the tumour samples.

**Allelic imbalance analyses**

The frequencies of tumours with AI at one or more loci at 9p21–22 or 10q26 were 67% and 71%, respectively. Representative samples are illustrated in Figure 2. When AI was present, it was usually seen at all informative loci along the same chromosomal region. There was no significant difference between the frequencies of imbalances in seminomas versus nonseminomas. Twenty-four out of 33 seminomas (73%) had AI at 9p21–22 versus 22 out of 35 nonseminomas (63%) ($P = 0.4$). At 10q26, 26 out of 33 seminomas (79%) versus 23 out of 35 nonseminomas (66%) had AI ($P = 0.3$).

**Comparison of methylation status and allelic imbalance**

Methylation status of the MGMT gene was compared to AI results for markers at 10q26, see Table 2. Among the 24 nonseminomas with methylated MGMT promoter, 17 showed AI at markers surrounding the gene (71%). In comparison, six out of 11 unmethylated nonseminomas had AI (55%). Among seminomas with methylated CpG sites, six out of eight samples had AI (75%), whereas in the group of unmethylated semi-

---

**Figure 2** Allelic imbalance at 9p21–22 and 10q26 loci in TGCT. Loss of alleles at microsatellite loci located to chromosome bands 9p21–22 and 10q26 is illustrated for two patients in the upper and lower panels, respectively. The genotype pattern is shown for constitutional and tumour DNA in each case. Peak heights reflecting the allele intensities in relative fluorescence units (y-axis) are given below each allele. The degree of allelic imbalance, calculated as $Q_{LOH}$, is shown for each locus. Note that constitutional homozygosity was found at D10S169.

---

Oncogene
nomas 20 out of 25 samples had AI (80%). No significant relation between methylation status and AI was found.

**MGMT protein expression**

Twenty of the TGCT were analysed by immunohistochemistry using an anti-MGMT antibody. Normal testicular tissues next to the tumours stained positive. Among the 13 TGCT not methylated in the MGMT promoter, seven had lost their MGMT protein expression (54%), whereas among the seven tumours methylated in the MGMT promoter, six had loss of expression (86%; $P = 0.18$, one-sided Fisher’s exact test). The methylated tumour with positive immunostaining was a nonseminoma expressing MGMT in its teratoma and yolk sac tumour components, but not in its embryonal carcinoma component.

**Discussion**

The present study is to our knowledge the first report on methylation of MGMT in testicular cancer. Aberrant methylation was observed in nearly half of the TGCT samples (32 out of 69), a frequency somewhat higher than previously reported for primary gliomas and colorectal carcinomas, two malignancies that have the highest known frequencies of hypermethylated MGMT (each approximately 40%) (Esteller et al., 1999). Indeed, this makes MGMT one of the most frequently altered genes identified in TGCT. Abnormal methylation of CpG sites in the promoter region can explain down-regulated MGMT expression in tumours with suppressed gene activity, with no evidence of large gene deletions or rearrangements. It is still unknown which of the CpG sites in the 5′-flanking region that need to be methylated, or how dense the methylation must be in order to silence gene expression. Studies have indicated that the specific methylation pattern as well as the degree of methylation may differ among samples (Costello et al., 1994; Qian et al., 1995). Detailed analyses of two cell lines have indicated hot spot regions from −249 to −103, and from +107 to +196 base pairs relative to the transcription start site (Qian and Brent, 1997). Further studies of more cell lines and xenografts have confirmed the association between methylation at certain CpG sites and lack of MGMT expression (Danam et al., 1999). One of the five CpG sites tested by Danam et al. (1999) is included in the regions screened in the present study, and the primer sites used are located within the designated hot spot region from +107 to +196. This region overlaps with an enhancer element of 59 base pairs previously reported to be necessary for efficient MGMT promoter function (Harris et al., 1994). Although less investigated than promoter methylation, methylation of some sites within the body of the gene may influence and be required for MGMT expression (Costello et al., 1994; Bearzatto et al., 2000). The fact that six out of seven methylated tumours of the present series had lost their MGMT protein expression support the view that the epigenetic changes detected cause gene silencing.

In addition to the relatively high overall MGMT methylation frequency in the present series, a significant difference in methylation was found between nonseminomas and seminomas. This cannot be due to differences in gene copy number since previous cytogenetic studies have reported a comparable low frequency of changes at chromosome arm 10q between the two subgroups. Indeed, a recent comparative genomic hybridization study of 33 TGCT, of which 30 are included in the present study, showed gain of 10q26 in only two seminomas and in two nonseminomas (Kraggerud et al., 2002). The present data of AI at 10q showed a much higher frequency of 10q changes than the cytogenetic data, possibly due to the different resolution levels of the two methods, but also by this methodological approach no clear difference between the two histological subgroups was found. It is therefore reasonable to assume that methylation of MGMT takes place during the transition from seminoma to nonseminoma, in line with the linear developmental model from carcinoma in situ, via seminoma into nonseminoma (Faulkner et al., 2000; Oosterhuis et al., 1989). A comparable situation has been reported for brain tumours (Nakamura et al., 2001). MGMT methylation was significantly more frequent in glioblastomas that had progressed from low-grade astrocytomas than in primary glioblastomas. However, the possibility remains that nonseminoma develops directly from a carcinoma in situ. Additional analysis of carcinoma in situ samples should provide further insights into the developmental model of TGCT.

The repair protein MGMT is able to remove alkyl groups from O6-guanine by an irreversible transfer to an internal cysteine residue. Increased repair may contribute to lower a cell’s sensitivity for, or development of resistance to, chemotherapy. Chemotherapy against malignant TGCT is based on cisplatin often in

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Methylation (n=8)</th>
<th>Unmethylated (n=25)</th>
<th>Methylation (n=24)</th>
<th>Unmethylated (n=11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{LOH}^{10q} \leq 0.84$ (=AI)</td>
<td>6</td>
<td>20</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>$Q_{LOH}^{10q} &gt; 0.84$</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

A tumour sample is denoted AI if one or more of the four markers along the chromosome arm had $Q_{LOH}^{10q} \leq 0.84$
combination with etoposide and bleomycin, and ifosfamide may be added when suboptimal response is observed. Ifosfamide is the only alkylating drug among these chemotherapeutic, but its cytotoxic effect does not seem to be influenced by MGMT activity (Preuss et al., 1996). It remains to be seen whether MGMT status is useful for identifying TGCT with specific chemo-sensitivity for certain alkylating agents. Interestingly, Esteller et al. (2000) reported an association between methylation of CpG sites in the MGMT gene, and responsiveness to carmustine for patients with gliomas.

Inactivation of growth-restricting genes may cause cells to proliferate more rapidly and therefore give them a selective advantage. Transcriptional repression of the CDKN2A gene by hypermethylation of 5’-CpG sites has specifically been seen in primary tumours and not in corresponding normal tissues (Merlo et al., 1995). The methylation assays used in the present study is limited to detection of CpG methylation status within short regions recognized by specific primers for unmethylated and methylated DNA templates. A strong inverse association has been reported between methylation detected by this assay and p16INK4a protein expression (Shim et al., 2000). Whether or not all CpG sites within each primer recognition site need to be methylated to obtain the PCR product may be sequence dependent, but methylation at CpG sites close to the elongation point is most likely a requirement. We have previously detected methylation within the CDKN2A gene in cases where subsequent sequence analysis confirmed complete CpG methylation (unpublished results). Thus, partial methylation will most likely not be detected by this approach and may explain the absence of CDKN2A methylation in the present series. The two other studies of CDKN2A gene methylation in TGCT samples both used restriction enzyme digestions to test for methylation in three CpG sites. In accordance with our findings, Heidenreich et al. (1998) did not find complete methylation in any of the tested samples, whereas Chaubert et al. (1997) reported methylation in 13 out of 26 samples without emphasizing whether or not the methylation patterns were complete.

In a cytogenetic survey of more than 3000 neoplasms, including 209 TGCT, net loss of chromosome arm 9p was deduced from each karyotype in about 30% of the TGCT (Mertens et al., 1997). The same frequency was found by comparative genomic hybridization among some of the tumours (n = 30) included in the present study (Kragerud et al., 2002). AI at 9p loci has previously been reported in two rather small TGCT series (Chaubert et al., 1997; Heidenreich et al., 1998). The high frequency of AI at 9p loci in the present series is in line with the findings of Heidenreich et al. (1998). Among 17 tumours, AI was found in 59% at one or more of five 9p markers, including two of the markers used in our study. In the other paper, no 9p losses were reported among 14 constitutionally heterozygous cases (Chaubert et al., 1997).

The present data suggest that alterations of MGMT contribute to development of the nonseminomatous histological subgroup of TGCT.

Materials and methods

Samples

Blood and primary tumour samples were obtained from 70 Norwegian patients with TGCT. The frozen tumour sample from each patient was sliced and three sections were stained with hematoxylin and eosin in order to estimate the fraction of intact neoplastic tissue. Across the sample set, an average of 75% neoplastic tissue was observed (range: 30 to 100%). The remaining tissue from each frozen sample was submitted to DNA extraction. Genomic DNA was isolated from blood and tumour samples in a 340A Nucleic Acid Extractor (Applied Biosystems, Foster City, CA, USA), applying standard phenol/chloroform extraction followed by ethanol precipitation. All frozen tumour samples were classified according to WHO recommendations (Mortioli and Sobin, 1976). Thirty-three of the samples were seminomas (25 were from patients with a pure seminoma and eight were from patients with mixed tumours (nonseminoma and seminoma components), 35 were nonseminomas, and two samples were carcinomas in situ (one represented carcinoma in situ nearby a seminoma, and the other was taken from testicular tissue surrounding a nonseminoma).

Methylation-specific PCR

DNA samples submitted to methylation analyses were modified according to the protocol of the CpGenomeTM DNA modification kit (InterGen, Boston, MA, USA). Two μg DNA was used as a starting amount and each treated sample was resuspended in 50 μl 1×TE, pH 7.5. DNA methylation status in the CpG islands of MGMT and CDKN2A were determined by subsequent PCR, using different primer sets specific for methylated and unmethylated CpG sites. Previously described primer sets were used for amplification of both MGMT fragments (Esteller et al., 1999) and CDKN2A fragments (Herman et al., 1996).

The MGMT fragments were amplified with 22.5 pmol of each primer, except for the anti-sense primer for unmethylated DNA where 5.6 pmol was used. The reaction mixtures contained 1.0 mM MgCl2 and 5 μl modified DNA solution as template. PCR was initiated by 7 min heating at 95°C, followed by 35 cycles of 95°C for 45 s, 58°C for 30 s and 72°C for 30 s, and ending by 7 min at 72°C. Human placental DNA (Sigma Chemical Co., St. Louis, MO, USA) treated in vitro with SssI methyltransferase (New England Biolabs Inc., Beverly, MA, USA) was the positive control for methylated templates, whereas DNA from normal testis tissue was used as the negative control. Methylation positive samples were visually scored as + or ++. Samples in the + + group were reproducibly amplified to an equal or higher intensity compared to the positive control of methylated DNA. The remaining positive samples, the + group, constituted reproducible methylated PCR products with less intensity than the positive control.

CDKN2A fragments were amplified with 25 pmol of each primer (Herman et al., 1996). Each reaction contained 1.5 mM MgCl2 and 2 μl modified DNA solution as template. The PCR program was as described above, except for the 64°C annealing temperature for unmethylated, and 63°C for
methylated DNA templates, and the 1 min long cycle steps. The human cell line T47D, earlier reported to harbour a methylated 5’-CpG island in CDKN2A (Herman et al., 1995; Hui et al., 2000), was the positive control. Again, DNA from normal testis tissue was the negative control.

All PCR reactions contained the following reagents: GeneAmp 1 × PCR Buffer II (Applied Biosystems), 200 µM dNTP (Amersham Pharmacia Biotech Products Inc., Piscataway, NJ, USA), and 2.5 U AmpliTaq DNA Polymerase (Applied Biosystems). PCR was performed in volumes of 50 µl using a GeneAmp PCR System 2400 (Applied Biosystems). The PCR products were analysed by running 7.5% PAGE, and stained with ethidium bromide before photographing using an UV transilluminator. All samples interpreted as methylated were confirmed by at least one photograph.

Allelic imbalance analyses

Eight polymorphic microsatellite loci, i.e. D10S1483, D10S1651, D10S1700 and D10S169, mapping to chromosome band 10q26, and D9S165, at 9p21 – 22, were analysed for AI (UCSC Genome Browser, April 2002 freeze, http://genome.ucsc.edu/). The samples were run on an ABI PRISM™ 310 Genetic Analyzer (Applied Biosystems). A semi-quantitative determination of the degree of AI, QLOH, was subsequently calculated from the measured peak heights by dividing the allele ratio in tumour DNA by the allele ratio in corresponding normal (blood) DNA. The cut-off level for AI scoring by this fluorescent primer protocol has previously been determined to QLOH ≤ 0.84 based on the variance of repeated analyses of 485 constitutional heterozygous genotypes at 20 different microsatellite loci (Skotheim et al., 2001b). A tumour sample is denoted AI if one or more of the four markers within the same chromosome region is registered with AI.

Immunohistochemistry

The MGMT protein expression was analysed by immunohistochemistry on formalin fixed and paraffin embedded tissues from twenty of the same tumours (seven methylated and 13 not methylated; nine seminomas and 11 nonseminomas). The biotin–streptavidin amplified system using the Optimax1 Automated Cell Staining System Plus (BioGenex, San Raman, CA, USA) was applied for immunohistochemistry of 4 µm tissue sections. After deparaffinization and rehydration, high temperature antigen retrieval was performed by three times 5 min microwaving at 900 W in 2 mCi citrate buffer, pH 6.0. Incubation with 1% hydrogen peroxide (H2O2) for 10 min was used to block the endogenous peroxidase activity before 30 min incubation in room temperature with the primary antibody (1:200 goat polyclonal anti-MGMT, sc-8825, Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA). The tissues were then incubated for 20 min with multilink biotinylated anti-immunoglobulins (1:30; BioGenex) and 20 min with streptavidin peroxidase (1:30; BioGenex). Further the sections were stained for 5 min with 0.05% of the peroxidase substrate 3,3-diaminobenzidine tetrahydrochloride (DAB) freshly prepared in 0.05 m Tris-HCl buffer at pH 7.6 containing 0.01% H2O2, before being counterstained with haematoxylin, dehydrated, and mounted. The immunostaining was nuclear, and cases with staining of more than 5% of the nuclei were considered positive.

Statistical analysis

Contingency tables were analysed by Fisher’s exact test. P-values less than 0.05 were considered statistically significant.

Acknowledgements

GE Lind and RI Skotheim are Research Fellows of Norwegian Cancer Society and Research Council of Norway respectively. The study was supported by additional grants from Norwegian Cancer Society (B Smith-Sorensen and RA Lothe).

References


Methylation of MGMT in testicular germ cell tumours
B Smith-Sørensen et al


The loss of NKX3.1 expression in testicular - and prostate - cancers is not caused by promoter hypermethylation

Molecular Cancer. 2005 Feb 3;4(1):8
The loss of NKX3.1 expression in testicular – and prostate – cancers is not caused by promoter hypermethylation

Guro E Lind¹, Rolf I Skotheim¹, Mario F Fraga², Vera M Abeler³, Rui Henrique⁴, Fahri Saatcioglu⁵, Manel Esteller², Manuel R Teixeira⁶ and Ragnhild A Lothe*¹,⁵

Abstract

Background: Recent studies have demonstrated that the NKX3.1 protein is commonly down-regulated in testicular germ cell tumors (TGCTs) and prostate carcinomas. The homeobox gene NKX3.1 maps to chromosome band 8p21, which is a region frequently lost in prostate cancer, but not in TGCT. Mutations have not been reported in the NKX3.1 sequence, and the gene is hypothesized to be epigenetically inactivated. In the present study we examined the methylation status of the NKX3.1 promoter in relevant primary tumors and cell lines: primary TGCTs (n = 55), intratubular germ cell neoplasias (n = 7), germ cell tumor cell lines (n = 3), primary prostate adenocarcinomas (n = 20), and prostate cancer cell lines (n = 3) by methylation-specific PCR and bisulphite sequencing.

Results and Conclusions: Down-regulation of NKX3.1 expression was generally not caused by promoter hypermethylation, which was only found in one TGCT. However, other epigenetic mechanisms, such as modulation of chromatin structure or modifications of histones, may explain the lack of NKX3.1 expression, which is seen in most TGCTs and prostate cancer specimens.

Background

The protein expression of the homeobox gene NK3 transcription factor related locus 1 (NKX3.1) is highly specific for the prostate and the testis [1-3], and is frequently lost in cancers of these two tissue types [1,4,5]. NKX3.1 is located in chromosome band 8p21 [2,6,7], a region that undergoes frequent allelic imbalance in prostatic intraepithelial neoplasia (PIN) and prostate carcinomas [8,9]. In mice, targeted disruption of Nkx3.1 leads to prostatic epithelial hyperplasia and dysplasia [10,11], and over-expression of exogenous NKX3.1 suppresses growth and tumorigenicity in human prostate carcinoma cell lines [12]. However, the expression levels and possible role for NKX3.1 during prostate cancer progression in humans is still being debated [13-15]. No gene mutations of NKX3.1 have been found [6], and NKX3.1 is therefore believed to be...
epigenetically inactivated in the cases with loss of protein expression [1,5,16]. Only one study has reported NKX3.1 protein expression in testicular germ cell tumors (TGCTs), however the series analyzed was large, including a total of more than 500 samples, and NKX3.1 was found absent in all embryonal carcinomas and present in only 15–20% of the seminomas as well as among the differentiated histological subtypes of germ cell tumors [5].

During the last decade, epigenetic changes in cancer have been frequently reported and are now recognized to be at least as common as genetic changes [17]. The best characterized epigenetic mechanism is DNA hypermethylation, in which cytosines located within selected CpG sites in the gene promoters become methylated, thereby inactivating gene expression. Several tumor suppressor genes are inactivated by such promoter hypermethylation in various cancer types [18,19]. In the present study we have performed methylation-specific PCR (MSP) and bisulphite sequencing of the \textit{NKX3.1} promoter in TGCTs and prostate adenocarcinomas to examine whether this mechanism may explain the commonly observed loss of NKX3.1 protein.

Results

Only one out of 54 TGCTs and none of the prostate adenocarcinomas (n = 20), intratubular germ cell neoplasias (n = 7), normal testis tissues (n = 4), or the cell lines (n = 6) displayed methylation when analyzed with MSP (Figure 1a). Bisulphite genomic sequencing of the tumors and cell lines showed that all cytosines at non-CpG sites were converted to thymine (Figure 1b). Only one sample demonstrated overall methylation in the \textit{NKX3.1} sequence, and this was the same sample that was positive for methylation from the MSP analysis. Interestingly, all the samples that were sequenced, including the normal blood, unmethylated cell lines, and primary tumors, displayed some extent of methylation (the majority below 25%) at the cytosine in CpG number 21 (base 1914762, +1 bp from transcription start). We detected a possible polymorphism in base 1914730 (+33 bp from transcription start and 15 bp upstream of the coding sequence). In previous sequences this site has been described as a guanine (Gene bank accession number NT_023666, and AF24770). In the cell lines, 5/6 contained adenosine in this position, but all except the germ cell tumor cell lines NCCIT and TERA2 were heterozygotes. In contrast, all 5 primary tumors sequenced were homozygous for the adenosine.

---

\textbf{Figure 1}

\textbf{Representative results of the methylation analyses of the \textit{NKX3.1} promoter.} (A) Methylation-specific PCR. A visible PCR product in \textit{Lanes U} indicates the presence of unmethylated alleles whereas a PCR product in \textit{Lanes M} indicates the presence of methylated alleles. The left panel illustrates the methylation status of selected TGCTs and the testicular cancer cell lines. Note the methylation of sample \# 2110. The right panel shows the unmethylated status of primary prostate cancers and prostate cancer cell lines. Abbreviations: NB, normal blood (positive control for unmethylated samples); IVD, \textit{in vitro} methylated DNA (positive control for methylated samples); neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product. (B) Bisulphite sequencing. The bisulphite sequence allows a positive display of 5-methyl cytosines in the gene promoter as unmethylated cytosines appear as thymines, while 5-methylcytosines appear as cytosines in the final sequence. The left chromatogram represents a part of the unmethylated \textit{NKX3.1} promoter in the germ cell tumor cell line TERA2, including 11 CpG sites marked by underlined letters. The right chromatogram represents the unmethylated prostate cancer cell line DU145. Both sequences have been generated by reversing the respective anti-sense sequences by use of the software "Chromas".
Discussion

We have previously reported that the protein expression of NKX3.1 is virtually ubiquitously lost in TGCTs [5]. This was done using a tissue microarray containing 510 testicular tissue samples. NKX3.1 expression is known for 25 of the TGCTs now analyzed for promoter hypermethylation and 22 showed complete absence of protein. The down regulation of NKX3.1 in TGCT has also been detected at the mRNA level, both by quantitative RT-PCR [5] and from an oligonucleotide microarray study including 20 of the present TGCTs (Skotheim et al., submitted). This simultaneous down regulation of both protein and mRNA levels of NKX3.1 is consistent with epigenetic regulation, which is further supported by the fact that mutations have not been detected in the NKX3.1 gene [6]. DNA promoter hypermethylation is the best-characterized epigenetic change in cancer, and can be associated with gene silencing. It was therefore of interest to analyze the methylation status of the NKX3.1 promoter in TGCT and prostate cancer samples. However, with the exception of a single TGCT, the NKX3.1 promoter was unmethylated in the samples analyzed. The methylated TGCT was classified as a yolk sac tumor, and has also been demonstrated to have promoter hypermethylation of several other genes that are generally unmethylated in TGCTs (Lind et al., unpublished). We therefore consider this sample not to be representative for the general TGCT epigenotype, nor for the general epigenetic profile of yolk sac tumors. Thus, we do not regard promoter hypermethylation as the general mechanism of NKX3.1 down-regulation neither in TGCT nor in prostate carcinomas.

We also studied cell lines since it can be argued that presence of normal cells as well as tumor heterogeneity may mask cancer specific methylation in primary tumors. LNCaP cells have previously been demonstrated to express NKX3.1, in contrast to PC-3 and DU-145, which do not express NKX3.1 since they lack a functional androgen receptor [2]. The lack of NKX3.1 expression in PC-3 and DU-145 cells is not due to methylation. This was also the case with the germ cell tumor cell lines. From Western analysis, the cell line NCCIT had strong expression of NKX3.1 whereas both TERA1 and TERA2 had no expression (data not shown).

The polymorphism in NKX3.1 that we detected 15 bases upstream of the coding sequence has to our knowledge not been described previously. It was identified by bisulphite sequencing of the cell lines and a subgroup of primary tumors, thus caution should be taken when concluding from these results, since regular sequencing analysis is the recommended approach for describing sequence changes. As the polymorphism is located in the promoter region of NKX3.1, it has no influence on the protein structure. However, it can still have a potential role in the transcriptional regulation of NKX3.1. A polymorphism in the coding sequence is also reported for NKX3.1 [6].

All samples analyzed with bisulphite sequencing, including cell lines expressing NKX3.1, as well as non-expressing cell lines, demonstrated some degree of methylation in the cytosine in CpG number 21. As this site-specific methylation included only one CpG site, it is unlikely that it will have any regulating effect on gene expression. However, considering its intriguing location immediate upstream of the transcription start point, this possibility should not be excluded. There is also the possibility that the apparent methylation could be due to a less efficient bisulphite conversion for this site. In general, the bisulphite sequencing results showed that all cytosines at non-CpG sites were converted to thymine (Figure 1b), but sequence-specific partial resistance to this conversion may lead to methylation artifacts, but only in rare cases, as has been reported previously [20].

Conclusions

In summary, these data show that the previously reported down-regulation of NKX3.1 in TGCTs and prostate carcinomas is not caused by promoter hypermethylation. Even though the NKX3.1 promoter is unmethylated, the simultaneous down-regulation of mRNA and protein levels in TGCTs and the absence of mutations still make other epigenetic mechanisms, such as modulation of chromatin structure or modifications of histones, possible explanations for loss of NKX3.1 expression in testicular- and prostate cancers.

Materials and Methods

Primary tumors and cell lines

Included in the present study are primary TGCTs (n = 55), intratubular germ cell neoplasias (also called carcinoma in situ; n = 7), normal testis tissue (n = 4), germ cell tumor cell lines (TERA1, TERA2, and NCCIT), prostate adenocarcinomas (n = 20), and prostate cancer cell lines (LNCaP, PC-3, and DU-145). The primary TGCTs include all histological subtypes: seminomas, embryonal carcinomas, teratomas, yolk sac tumors, and one choriocarcinoma, classified according to the WHO's recommendations [21] by a germ cell tumor reference pathologist using light microscopic examination of hematoxylin and eosin stained tissue sections. From our previous comparative genome hybridization analysis, about half of the TGCTs had a low-level copy number gain at chromosome 8, but only rarely 8p deletions [22]. Primary prostate adenocarcinomas obtained from radical prostatectomy specimens were graded according to the Gleason grading system [23] using routinely stained tissue sections. The median
Gleason score of prostate adenocarcinomas was 7 (range: 4 – 8). The prostate carcinomas were all of pTNM stage 2 and 3, and included 10 samples with 8p deletions (among other cytogenetic aberrations), 3 samples with copy number changes not involving the 8p region, and 7 samples with no copy number changes (Ribeiro et al., submitted).

Methylation-specific PCR
The DNA samples were initially bisulphite modified [24,25], which converts unmethylated but not methylated cytosines to uracil. All samples were subsequently submitted to MSP analysis [26] using PCR primers specific to methylated and unmethylated sequences: NKX3.1 unmethylated sequence, sense: 5’GCAAAGTCAAGTGGTGTGGTT3’, antisense: 5’CTACACACATCCACAAATATC3’; methylated sequence, sense: 5’AAAGTGAAACCGTGCCGGTGC3’, antisense: 5’ACGCGGCCGTCCCGCAAAATAT3’ (MedProbe AS, Oslo, Norway) represented the unmethylated positive control and the methylated positive control, respectively. Water, replacing bisulphite treated template, was the negative control in both reactions.

Bisulphite sequencing
Bisulphite sequencing allows a positive display of 5-methyl cytosines in the gene promoter after bisulphite modification as unmethylated cytosines appear as thymines, while 5-methylcytosines appear as cytosines in the final sequence [27]. A subset of the samples (n = 11) were bisulphite sequenced, including all 6 cell lines, 3 TGCTs, and 2 prostate adenocarcinomas. Additionally, NB and IVD were bisulphite sequenced as positive controls for unmethylated and methylated sequence, respectively. The NKX3.1 bisulphite sequence fragment (Gene bank accession number NT_023666 (minus strand), bases 1914526 to 1914961) was 436 bp long and covered 52 CpG sites in the promoter and first exon of the gene. We designed bisulphite sequencing primers (MedProbe) with the following sequences; sense: 5’ATTGGGGAAGGAGAGGGAATTG3’, antisense: 5’CCCTAATCTCTAACTCTAACTCC3’. The Mg2+ content of the reaction was 1.3 mM, the enzyme used was HotStarTaq™ DNA polymerase (QIAGEN Inc., Valencia, CA, USA), and the annealing temperature 52°C. The PCR fragments were eluted from a 2% agarose gel (BioRad Laboratories Inc, CA, USA) containing ethidium bromide, by the MinElute™gel extraction kit (QIAGEN), and sequenced with the dGTP BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) in an ABI Prism 377 Sequencer (Applied Biosystems). The bisulphite sequencing results were scored according to Melki et al. where the amount of methylcytosine of each CpG dinucleotide is quantified by comparing the peak height of the cytosine signal with the sum of the cytosine and thymine peak height signals [28].

Authors’ contributions
GEL performed the experimental analyses and statistics, interpreted the results, and drafted the manuscript. RIS did an independent scoring of the results, and contributed to manuscript preparation. MEF designed the primers used for the MSP and bisulphite treatment and contributed to manuscript preparation. VMA and RH were reference pathologists for the testicular cancer tissues and prostate tissues, respectively. FS was responsible for the Western Blot studies of the cell lines and participated in the writing of the manuscript. Parts of this work were done in the lab of ME who also contributed with scientific discussions. MRT provided the relevant selected series of primary prostate carcinomas with known genetic profiles, and contributed to manuscript preparation. RAL conceived the study, was responsible for its design and coordination, and contributed in the evaluation of the results and in preparation of the manuscript.

Acknowledgements
The study was funded by grants from the Norwegian Cancer Society supporting GEL as a Research Fellow and RIS as a Post-Doctoral Fellow (A95068 RAL). The authors are grateful to Gunnar Brunborg at the Norwegian Institute of Public Health, Oslo, Norway for providing the normal testis DNA from organ donors.

References
7. Swalwell JI, Vocke CD, Yang Y, Walker JR, Grouse L, Myers SH, Gillespie JW, Bostwick DG, Duray PH, Linehan WM, Emmert-Buck MR: Determination of a minimal deletion interval on chro-
Epigenetically deregulated genes novel in testicular cancer development include

homeobox genes and \textit{SCGB3A1 (HIN-1)}

Submitted manuscript
Epigenetically deregulated genes novel in testicular cancer development include homeobox genes and SCGB3A1 (HIN-1)

GE Lind\textsuperscript{1}, RI Skotheim\textsuperscript{1}, MF Fraga\textsuperscript{2}, VM Abeler\textsuperscript{3}, M Esteller\textsuperscript{2}, RA Lothe*\textsuperscript{1,4}

\textsuperscript{1}Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital, 0310 Oslo, Norway; \textsuperscript{2}Cancer Epigenetics Laboratory, the Spanish National Cancer Centre (CNIO), Madrid, Spain; \textsuperscript{3}Department of Pathology, The Norwegian Radium Hospital, 0310 Oslo, Norway; \textsuperscript{4}Department of Molecular Biosciences, Institute of Biology, University of Oslo, Oslo, Norway.

Testicular germ cell tumors (TGCTs) are classified as seminomas and nonseminomas. The latter histological group comprises embryonal carcinomas, yolk sac tumors, choriocarcinomas, and teratomas. The embryonal and extra-embryonal differentiation lineages among these subtypes parallel normal development, and inactivation by promoter hypermethylation of genes important in embryogenesis might contribute to testicular tumorigenesis. We therefore analyzed the following homeobox genes DLX6, EMX2, HOXA9, HOXB5, MSX1, and MSX2, as well as four other selected genes, CDH13, RASSF1A, RUNX3, and SCGB3A1 (alias HIN-1) in 7 intratubular germ cell neoplasias and 55 primary TGCTs representing all histological subtypes. In addition, by a discovery based global approach comparing cDNA microarray expression profiles of germ cell tumor cell lines before and after treatment with 5-aza-2'-deoxycytidine, we identified a gene list of potential epigenetic targets, from which CGGBP1, CGRRF1, SMARCC2, SORBS1, and XPA were analyzed further. Overall, the nonseminomas were by far more often methylated than seminomas (median 0.10, and median 0.0000 respectively, \(P < 0.001\)). The three most frequently methylated genes were SCGB3A (HIN-1), methylated in 54\% of the nonseminomas, RASSF1A (29\%), and HOXA9 (26\%). CDH13 and HOXB5 demonstrated methylation at low frequencies, and EMX2, MSX1, RUNX3, and SORBS1 only rarely. In summary, the methylation frequency variations along the differentiation lineages of the testis tumor model are comparable with the methylation levels seen at various developmental stages in the early mouse embryo, initially caused by epigenetic reprogramming. HOXA9 and SCGB3A1 (HIN-1) are novel epigenetically regulated targets in testicular tumorigenesis.

Introduction

Testicular cancer is the most common malignant disease among young Caucasian men, and the incidence has increased the last decades (Bergström \textit{et al}., 1996); (Møller, 2001; McGlynn \textit{et al}., 2003). Most of the tumors are of germ cell origin and are named testicular germ cell tumors (TGCTs), accordingly. These tumors are divided into two main histological classes, seminomas and nonseminomas (Mostofi and Sesterhenn, 1998; Skotheim and Lothe, 2003), both suggested to develop from the precursor stage intratubular germ cell neoplasia (also called carcinoma \textit{in situ}) (Skakkebaek \textit{et al}., 1987). The undifferentiated seminomas morphologically resemble the intratubular germ cell neoplasia, whereas the nonseminomas include several histological subtypes along complete differentiation lineages. Cells of the undifferentiated pluripotent embryonal carcinoma may differentiate into teratomas, containing tissues from all three germ layers, or they may differentiate along extra-embryonal lineages into yolk sac tumors or choriocarcinomas (Skakkebaek \textit{et al}., 1987; Damjanov, 1990; Mostofi and Sesterhenn, 1998).

Testicular germ cell tumorigenesis may in many ways be looked upon as a genetic and epigenetic caricature of early embryogenesis (Andrews, 2002). Embryonal carcinoma cells share several of the features of the inner cell mass in the human blastocysts and their derived embryonic stem cells (Henderson \textit{et al}., 2002), which is underlined by the similarities in gene expression (Sperger \textit{et al}., 2003). Some of the embryonal carcinoma cell lines are able to differentiate in culture in

*Correspondence: RA Lothe; E-mail: rlothe@radium.no
response to specific agents like retinoic acid (Andrews et al., 1984), and they have the ability to divide when transplanted into the blastocyst (Brinster, 1974). Recently, some of the most common genomic changes found in TGCT were also identified in cultures of embryonal stem cells (Draper et al., 2004). An epigenetic link between testicular tumorigenesis and embryogenesis is implied by a similar expression pattern of imprinted genes in normal embryonic development and TGCT (van Gurp et al., 1994).

Promoter hypermethylation and associated gene silencing is widely recognized as an inactivating mechanism of tumor suppressor genes in most types of cancer (Esteller et al., 2001), and to date, numerous genes have been found to undergo hypermethylation. So far, about 25 different genes commonly methylated in various cancer types have been analyzed in TGCT, but only a handful have demonstrated methylation in more than a small subset of cases (Koul et al., 2002; Smith-Sørensen et al., 2002; Honorio et al., 2003). Among the most frequently methylated genes are MGMT, RASSF1A, APC, HIC1, and BRCA1 (Koul et al., 2002; Smith-Sørensen et al., 2002; Honorio et al., 2003; Koul et al., 2004). Other candidate target genes, like CDKN2A (alias p16INK4a), commonly methylated in a variety of cancers (Herman et al., 1995; Esteller et al., 2001), have been found unmethylated in TGCT (Smith-Sørensen et al., 2002).

The homeobox gene family encodes transcription factors (Levine and Hoey, 1988) regulating morphogenesis and cell differentiation during embryogenesis, and their functions are critical in specifying cell identity and positioning in the developing embryo (Krumlauf, 1994). These genes contain a first identified 180 bp stretch DNA, encoding a 60 amino acid homeodomain, and were first conserved 180 bp stretch DNA, encoding a 60 amino acid homeodomain, and were first identified in the fruit fly Drosophila melanogaster (Lewis, 1978). The homeodomain, organized in a helix-turn-helix conformation, is responsible for recognition and sequence-specific DNA binding, and thereby the homeodomain protein activates or represses expression of batteries of downstream target genes (Boersma et al., 1999). In addition to being important during embryogenesis, several homeobox genes seem to play a role in the maintenance of homeostasis, as they are also expressed in some normal adult tissues, including the testis (Takahashi et al., 2004). Altered expression of homeobox genes has long been associated with different types of cancer (Abate-Shen, 2002).

To our knowledge promoter methylation of homeobox genes in TGCT remain unknown. In the present study we analyzed primary TGCTs, precursor lesions: intratubular germ cell neoplasias, germ cell tumor cell lines, and normal testicular tissues for promoter hypermethylation of six such genes, empty spiracles homolog 2 (EMX2), distal-less homeo box 6 (DLX6), homeo box A9 (HOXA9), homeo box B5 (HOXB5), msh homeo box homolog 1 (Drosophila) (MSX1), msh homeo box homolog 2 (Drosophila) (MSX2). We also analyzed four additional candidate genes; H-cadherin (CDH13), runt-related transcription factor 3 (RUNX3), secretoglobin, family 3A, member 1 (SCGB3A1; alias HIN-1), and Ras association (RalGDS/AF-6) domain family 1 (RASSF1A). Finally, to identify a list of potentially epigenetically regulated genes in TGCTs, we used a cDNA microarray and compared the gene expression pattern in germ cell tumor cell lines with and without an induced global demethylation by use of 5-aza-2'-deoxycytidine. The candidates on the list of re-expressed genes were examined for down-regulation in an existing DNA microarray expression data set of primary TGCT, and the commonly altered genes were submitted to downstream methylation analyses. Figure 1 illustrates the two combined approaches: candidate gene driven- and global discovery based- analyses.

Results

Methylation status of candidate genes

The results of the MSP analyses of 55 TGCTs and 7 intratubular germ cell neoplasias are summarized in Figure 2 and representative raw data can be seen in Figure 3. Promoter methylation was significantly more frequent in nonseminomas than in seminomas for SCGB3A1 (19/35 versus 0/20, P = 0.001). RASSF1A (10/35 versus 0/19, P = 0.010), and HOXA9 (9/35 versus 0/20, P = 0.019). Also HOXB5, EMX2, MSX1, and RUNX3, were methylated only in nonseminomas, although at a lower frequency (Figure 2a). In general, the nonseminomas demonstrated a 26-fold higher mean of methylation frequency across the ten initially analyzed genes than did the seminomas (nonseminomas: median 0.1000, mean 0.1465; seminomas: median 0.0000, mean 0.0056, P < 0.001). Also, within the nonseminoma group, statistically significant differences in the mean frequencies of methylated genes were demonstrated as the embryonal carcinomas contained less methylation than the rest of the nonseminomas.
**Candidate gene approach**

*Selected genes*

- **a)** Developmental genes: *EMX2, DLX6, HOXA9, HOXB5, MSX1*, and *MSX2*
- **b)** Methylated in cancer: *CDH13, RUNX3*, and *SCGB3A1*
- **c)** Methylated in TGCT: *RASSF1A*

**Discovery based approach**

*Expression profiling by cDNA microarrays*

- **a)** Cell lines cultured with and without 5-aza-2’deoxycytidine
- **b)** Comparison with down-regulated expression in primary TGCT

**Genes submitted to methylation analysis, n = 10**

**Gene expression patterns of 10 selected genes in human testis tissue samples**

- **a)** 99 upregulated genes after treatment
- **b)** 28 of the 99 were downregulated in TGCT

**Genes selected for methylation analysis, n = 5**

- *CGGBP1, CGRRF1, SMARCC2, SORBS1, XPA*

**Methylated genes in TGCT:**

*SCGB3A1, RASSF1A, HOXA9, CDH13, HOXB5, SORBS1, EMX2, MSX1, RUNX3, XPA*

**Figure 1. The study design.** Two complimentary approaches for identifying new target genes methylated in TGCTs. Note that the gene expression panel to the right has been clustered for visualization. Abbreviations: TGCT, testicular germ cell tumors; ITGCN, intratubular germ cell neoplasia. *The gene CGRRF1 was upregulated less than two times in the cell lines after 5-aza-2 deoxycytidine treatment, but was still selected for methylation analysis based on its’ several times down-regulation in all primary tumors.*
Embryonal carcinomas: median 0.0000; remaining nonseminomas: median 0.2000, \(P = 0.001\) and the teratomas seemed to display more methylation than did the rest of the nonseminoma group (teratomas: median 0.2; remaining nonseminomas: median 0.1, \(P = 0.07\))(Figure 2b). The primary tumors analyzed in the present study included only one choriocarcinoma, which was methylated in 3/8 genes (not determined for \(DLX6\) and \(HOXB5\)). One of the TGCT samples, a yolk sac tumor, displayed unusually high levels of methylation as it was simultaneously methylated in 7/10 genes. This was the only sample methylated for \(EMX2\), \(MSX1\), and \(RUNX3\). \(CDH13\) was the only gene with methylation in a seminoma sample (1/17). \(DLX6\) and \(MSX2\) were unmethylated for all samples analyzed. Methylation status was also tested in four normal testicular samples in the following genes: \(CDH13\), \(HOXA9\), \(MSX1\), \(MSX2\), \(RUNX3\), and \(SCGB3A1\), which were all unmethylated.

Identicalation of a new gene list of potentially epigenetically regulated genes in TGCT

Prior to cDNA microarray analyses, the three testicular cancer cell lines were cultured in medium with and without 5-aza-2'-deoxycytidine. This led to the appearance of a new gene list of potentially epigenetically regulated genes in TGCT.
of unmethylated MSP bands for the cell lines originally methylated for SCGB3A1 (Figure 5). From the cDNA microarray assays, we generated a list of 99 annotated candidate genes for methylation (upregulated two or more times upon treatment with 5-aza-2'-deoxycytidine in at least one of the cell lines). Among the 99 genes, 28 were down-regulated (Figure 1) as assessed by oligo microarrays in 23 primary TGCTs relative to normal testis samples; IVD, in vitro methylated DNA (positive control for methylated samples); neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product. The gel panels are each a merge of results from up to three separate rounds of MSP and subsequent electrophoresis.

Discussion

This study emphasizes the importance of epigenetic regulation of genes during development of TGCT. We have shown promoter hypermethylation of genes novel in TGCT, including homeobox genes, a family with important functions during normal development. It has previously been shown that the nonseminoma group of TGCT displays considerably more methylation than do the seminomas (Smiraglia et al., 2002; Smith-Sørensen et al., 2002). Indeed, this overall characteristic is confirmed in the present study. Additionally, by use of tissue samples containing only a single histological subgroup, we demonstrate specific methylation patterns also among the nonseminoma subtypes. These histological subtype-specific epigenotypes seem to parallel the epigenetic reprogramming of the early embryonic development.

In mice, methylation reprogramming occurs in preimplantation embryos (Reik et al., 2001). Both the parental and the maternal genomes are demethylated immediately after fertilization, resulting in hypomethylation of the genomes at the morula stage (Mayer et al., 2000). The DNA of blastocysts is thus relatively undermethylated, but after implantation, the genomes of the cells in the inner cell mass become de novo methylated. This genome-wide de novo methylation accompanies the loss of totipotency and is the first differentiation event in mammalian embryos (Dean et al., 2001). The embryonal carcinoma stage in TGCT share several features with cells in the inner cell mass, including pluripotency and gene expression similarities (Henderson et al., 2002; Sperger et al., 2003). In the present study embryonal carcinomas display the lowest mean methylation frequency among all the nonseminoma subgroups. Although most of the genes analyzed here for promoter methylation were unmethylated in the embryonal carcinomas, a few exceptions, like SCGB3A1, were methylated (Figure 2b), all in concordance with the embryonic development. As the gene expression of both the de novo methyltransferases DNMT3A and B are upregulated in embryonal carcinomas (Skotheim et al., 2005), promoter methylation of specific genes is expected. However, a threshold level is reached upon methylation of several genes, and the embryonal carcinoma might initiate differentiation into other histological subtypes and thereby loose its pluripotent capabilities.

As the mouse blastocyst develops into embryonic and extra-embryonic lineages, the genome is remethylated to different extents leading to more methylation in the embryonic lineages than the extra-embryonic ones (Reik et al., 2001). In testicular germ cell tumorigenesis, teratomas represent fully differentiated somatic tissues of various types. By group wise comparison, this histological...
Figure 4. Bisulphite sequencing results of HOXA9 and SCGB3A1 in representative samples. Several clones were sequenced from each sample, revealing the methylation status of individual CpG sites in a) HOXA9 and b) SCGB3A1. The upper part of each panel is a schematic presentation of the CpG island in the area of transcription start amplified by the bisulphite sequencing primers. The transcription start site is represented by +1 and the vertical bars indicate the location of individual CpG sites. The arrows indicate the location of the methylation specific-PCR primers. Filled circles, methylated CpGs; open circles, unmethylated CpGs. The column in the right side of each panel lists the methylation status for individual samples from MSP analyses. U, unmethylated; M, methylated.

subgroup showed the highest frequency of methylation among the nonseminoma subtypes, although this was not evident at the single gene level. We also detected relatively high methylation frequencies among yolk sac tumors, which are sharing morphological characteristics with the extra-embryonic yolk sac. One of the yolk sac tumors displayed exceptionally frequent promoter methylation, and we believe that the methylation machinery in this single sample is somehow overactive, resulting in a methylation pattern that is not representative for yolk sac tumors as a group. If we exclude this yolk sac tumor sample, RASSF1A is the only gene displaying more methylation in yolk sac tumors than in teratomas, an observation in accordance with a previous report (Koul et al., 2002). Frequent RASSF1A methylation in the nonseminoma subtype has also been reported by Honorio and co-workers (Honorio et al., 2003).
Figure 5. Demethylation of the SCGB3A1 promoter after 5-aza-2'-deoxycytidine treatment. A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. a) All three germ cell tumor cell lines analyzed displayed initial SCGB3A1 promoter hypermethylation. b) After 5-aza-2'-deoxycytidine treatment of the cell lines, unmethylated gel bands appeared in addition to the methylated ones, indicating demethylation of SCGB3A (although not complete). Abbreviations: NB, normal blood (positive control for unmethylated samples); IVD, in vitro methylated DNA (positive control for methylated samples); neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product.

We report for the first time promoter hypermethylation of SCGB3A1 (HIN-1) in TGCT. SCGB3A1 was initially characterized based on its differential expression between malignant and normal breast tissues (Krop et al., 2001). It was named "high in normal-1" (HIN-1) as the gene expression was significantly down regulated in the majority of breast carcinomas. SCGB3A1, which is the approved gene symbol for HIN-1 according to the HUGO Gene Nomenclature Committee, turned out to have a hypermethylated promoter in a significant number of primary breast tumors analyzed (Krop et al., 2001). Recently, the gene was also reported down-regulated in non-small cell lung cancers (Marchetti et al., 2004), primary nasopharyngeal carcinomas (Wong et al., 2003), and a variety of other cancer types (Shigematsu et al., 2005), but the exact function of the protein remains unknown. Both the initial study of SCGB3A1 (Krop et al., 2001) as well as subsequent studies (Wong et al., 2003; Shigematsu et al., 2005) have documented a strong association between promoter hypermethylation assessed by MSP and reduced gene expression. In the present study we have used the same MSP primer set and identified a high methylation frequency of SCGB3A1. However, by bisulphite sequencing we demonstrate that some of the methylated samples from the MSP analyses might be false positives, as two of the "methylated" samples submitted to sequencing analyses only contained few and scattered 5-methylcytosines in the 5' end of the gene. False positives resulting in an overestimation of methylation frequencies is a common problem using MSP analyses since incompletely converted DNA from bisulphite treatment can be coamplified with the methylated MSP reaction (Rand et al., 2002). However, incompletely converted DNA can only in part explain the present results. From Figure 4 we see that TGCT 307, which was scored as methylated from MSP analyses, is basically unmethylated when SCGB3A1 is bisulphite sequenced. If this was caused by incomplete base conversion, unmethylated cytosines would be expected to be present also when sequencing other genes from the same sample. However, all HOXA9 clones bisulphite sequenced from the same sample were methylated. Even though the methylation frequency presented here might be slightly overestimated, the bisulphite sequencing results simultaneously confirm that SCGB31 is a target for aberrant hypermethylation in TGCTs.

Methylation of CDH13, a member of the cadherin family, encoding H-cadherin, a cell surface glycoprotein involved in cell adhesion, has been reported to associate with poorly differentiated colorectal cancers (Hibi et al., 2004) as well as colorectal adenomas (Toyooka et al., 2002). Interestingly, the CDH13 is unmethylated in the highly differentiated teratoma group, and methylated among some samples belonging to the undifferentiated subgroups (two embryonal carcinoma and one seminoma). Across all the genes analyzed in the present study, teratomas demonstrated to be the most frequently methylated histological subgroup, whereas embryonal carcinomas displayed the least methylation. One might speculate that inactivation of CDH13 provide the undifferentiated tumor cells with a selective growth advantage due to reduced cell adhesion. Once fully differentiated, the methylation does not need to be sustained.

Although homeobox genes have been studied extensively, and their expression profiles determined in a number of human tissues (Care et al., 1996; Chariot and Castronovo, 1996; Cillo et al., 2001; Takahashi et al., 2004), little is known regarding the establishment of their methylation patterns, both in normal tissues and cancers. However, their role as crucial transcription factors in embryogenesis makes them candidates for involvement also in tumor development. In the present study we demonstrate a high frequency of HOXA9 methylation in nonseminomatous TGCTs, confirmed by bisulphite sequencing. The distribution of methylation frequencies within the histological subgroups was comparable with the non-
Epigenetic changes in testicular germ cell tumors

Homeobox genes analyzed (Figure 2b). Three additional homeobox genes, EMX2, HOXB5, and MSX1, also displayed methylation, but infrequent and only in nonseminomas. The hypermethylated promoters of specific homeobox genes indicate that developmental genes play an important role in the testicular tumorigenesis, and further supports its mimicry of the early embryogenesis. However, the homeobox genes’ functional relevance to the neoplastic phenotype remains to be elucidated.

We also searched for new candidate genes for methylation in TGCT using a global genome approach in cell lines, which were cultured with and without 5-aza-2'-deoxycytidine, a compound that in effect removes the DNA methylation. From the cDNA microarray data, we generated a list of potentially methylated genes in TGCTs based on the upregulated gene expression after removal of DNA methylation. This gene list was compared with expression of the same genes in primary testicular tumors (from DNA microarray analyses, data not shown), and from this we selected five down regulated genes with a CpG island located in the 5' end and encoding proteins with functions that could play a potential role in tumorigenesis. One of the candidates, XPA, has a crucial role in the nucleotide excision repair pathway. The exceptional sensitivity of TGCTs of adolescents and adults to chemotherapy, in particular to cisplatin, has been attributed to low levels of XPA (Koberle et al., 1999). However, a recent report shows that although testis tumor cell lines have low levels of XPA protein, the XPA mRNA levels are not reduced (Welsh et al., 2004). This is in accordance with the present finding of the XPA promoter as rarely methylated in TGCT.

CGRRF1 is encoding a cell growth regulator protein with a ring finger domain. This domain is a binding motif for the ubiquitin conjugating enzyme, and the ring finger proteins function as ubiquitin-protein ligases for heterologous substrates as well as for ring finger proteins themselves (Joazeiro and Weissman, 2000). Mutations in other ring finger encoding genes, like BRCA1, are seen in several types of human cancer like familial breast - and ovarian - cancer (Miki et al., 1994). BRCA1 is additionally silenced by promoter hypermethylation in the same cancer types (Esteller et al., 2001) as well as in TGCT (Koul et al., 2002). However, the ring finger candidate gene CGRRF1 was unmethylated in all the primary TGCTs. The remaining three candidate genes for methylation in testicular cancer were also unmethylated despite the significantly down regulated gene expression in TGCT, the reactivated expression in 5-aza-2'-deoxycytidine treated cell lines, the CpG islands in the promoter region, and their potential cancer related protein function. The low mRNA levels of these genes must therefore be caused by some other mechanism than promoter methylation.

We have identified several novel gene targets in TGCT development and suggest that epigenetic inactivation of many key genes in normal development also have an important role in TGCT as shown by epigenetic inactivation of several homeobox genes. Our results also indicate that methylation changes in testicular tumorigenesis mimic the epigenetic reprogramming during early embryogenesis.

Material and methods

Tissue samples and cell cultures

DNA from 55 primary TGCTs and 7 intratubular germ cell neoplasias from 61 patients was submitted to methylation analyses, along with 3 germ cell tumor cell lines (TERA1, TERA2, and NCCIT), and 4 normal testis tissue samples. Three of the normal samples were from organ donors with no known history of cancer and the last sample from normal tissue adjacent to cancer. All tumor samples have been classified according to the WHO's recommendations (Mostofi and Sesterhenn, 1998). Twenty of the TGCT samples were seminomas, and 35 were nonseminomas. The nonseminomas consisted of 16 embryonal carcinomas, 9 teratomas, 6 yolk sac tumors, 1 choriocarcinoma, and 3 tumors with mixed nonseminoma-components. The cell lines were all of embryonal carcinoma origin (Fogh, 1978; Andrews et al., 1984; Teshima et al., 1988). TERA1 and TERA2 were cultured in McCoy's 5a medium (Sigma; Chemical Co., St. Louis, MO) with 15% fetal bovine serum (GIBCO™, Paisley, Scotland, UK), 1.5 mM L-glutamine (GIBCO), 100 U/ml penicillin and 100 mg/ml streptomycin (GIBCO), and 0.25 mg/ml fungizone amphotericin B (GIBCO), whereas NCCIT was cultured in RPMI1640 (GIBCO) with the same additions as to the McCoy's 5a medium, except for the L-glutamine. The cell lines were cultured in parallel with and without 10 µM 5-aza-2'-deoxycytidine (Sigma) in the medium for 72 hours.

Methylation-specific PCR (MSP) and bisulphite sequencing

Promoter methylation was initially studied in CDH13, DLX6, EMX2, HOXA9, HOXB5,
Table 1. PCR primers used for MSP and bisulphite sequencing. Abbreviations: MSP, methylation specific-PCR; BS, bisulphite sequencing; M, methylated-specific primers; U, unmethylated-specific primers; Frg. Size, fragment size; An. Temp, annealing temperature *The annealing temperature is given in degrees celsius. ¤Fragment location lists the start and end point base pair of each fragment relative to the transcription start point provided by NCBI (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi.)

<table>
<thead>
<tr>
<th>Primer set</th>
<th>Sense primer</th>
<th>Antisense primer</th>
<th>Frg. Size</th>
<th>An. Temp*</th>
<th>Fragment location¢</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDH13 MSP-M</td>
<td>TCGCGGTTGTTTTCGCGTTT</td>
<td>GACGTTTATCATCATACAGCGCG</td>
<td>243</td>
<td>58</td>
<td>-146 to +97</td>
</tr>
<tr>
<td>CDH13 MSP-U</td>
<td>TCGCGGTTGTTTTCGCGTTT</td>
<td>AAGCTTATCATCATACAGCGCG</td>
<td>243</td>
<td>58</td>
<td>-146 to +97</td>
</tr>
<tr>
<td>CGGBP1 MSP-M</td>
<td>CGCGGTTGTTTTCGCGTTT</td>
<td>AAAATCGGCTGCTACCGCGC</td>
<td>130</td>
<td>65</td>
<td>-98 to +32</td>
</tr>
<tr>
<td>CGGBP1 MSP-U</td>
<td>CGCGGTTGTTTTCGCGTTT</td>
<td>AAAATCGGCTGCTACCGCGC</td>
<td>130</td>
<td>65</td>
<td>-98 to +32</td>
</tr>
<tr>
<td>CGRRF1 MSP-M</td>
<td>CGCGGTTGTTTTCGCGTTT</td>
<td>CCGCGCGACGCTACGCTACAG</td>
<td>157</td>
<td>62</td>
<td>-193 to -36</td>
</tr>
<tr>
<td>CGRRF1 MSP-U</td>
<td>CGCGGTTGTTTTCGCGTTT</td>
<td>ACGCGCGACGCTACGCTACAG</td>
<td>157</td>
<td>62</td>
<td>-193 to -36</td>
</tr>
<tr>
<td>DLX6 MSP-M</td>
<td>CTCGCGGTTGTTTTCGCGTTT</td>
<td>GCCCGGACGCTGCTACGCTACAG</td>
<td>131</td>
<td>65</td>
<td>-10 to +121</td>
</tr>
<tr>
<td>DLX6 MSP-U</td>
<td>CTCGCGGTTGTTTTCGCGTTT</td>
<td>GCCCGGACGCTGCTACGCTACAG</td>
<td>131</td>
<td>65</td>
<td>-10 to +121</td>
</tr>
<tr>
<td>EMX2 MSP-M</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>ACGGCTACGCTGCTACGCTACAG</td>
<td>127</td>
<td>66</td>
<td>-65 to +62</td>
</tr>
<tr>
<td>EMX2 MSP-U</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>ACGGCTACGCTGCTACGCTACAG</td>
<td>127</td>
<td>66</td>
<td>-65 to +62</td>
</tr>
<tr>
<td>HOXA9 MSP-M</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CCATACGCGCTGCTACGCTACAG</td>
<td>139</td>
<td>66</td>
<td>-97 to +69</td>
</tr>
<tr>
<td>HOXA9 MSP-U</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CCATACGCGCTGCTACGCTACAG</td>
<td>139</td>
<td>66</td>
<td>-97 to +69</td>
</tr>
<tr>
<td>MSX1 MSP-M</td>
<td>CCACGCGGTTGTTTTCGCGTTT</td>
<td>AAATCCGTCGCTACCGCGC</td>
<td>134</td>
<td>62</td>
<td>-82 to +19</td>
</tr>
<tr>
<td>MSX1 MSP-U</td>
<td>CCACGCGGTTGTTTTCGCGTTT</td>
<td>AAATCCGTCGCTACCGCGC</td>
<td>134</td>
<td>62</td>
<td>-82 to +19</td>
</tr>
<tr>
<td>MSX2 MSP-M</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CTACGCGCTGCTACGCTACAG</td>
<td>113</td>
<td>66</td>
<td>+6 to +119</td>
</tr>
<tr>
<td>MSX2 MSP-U</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CTACGCGCTGCTACGCTACAG</td>
<td>113</td>
<td>66</td>
<td>+6 to +119</td>
</tr>
<tr>
<td>RASSF1A MSP-M</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CCGCGGACGCTGCTACGCTACAG</td>
<td>229</td>
<td>62</td>
<td>+218 to +447</td>
</tr>
<tr>
<td>RASSF1A MSP-U</td>
<td>GGTGTTGTTGTTTTCGCGTTT</td>
<td>CCGCGGACGCTGCTACGCTACAG</td>
<td>229</td>
<td>62</td>
<td>+218 to +447</td>
</tr>
<tr>
<td>RUNX3 MSP-M</td>
<td>TTACCGCGGTTGTTTTCGCGTTT</td>
<td>AAAAAATCACGTACTTTTATTAACCAATCG</td>
<td>113</td>
<td>53</td>
<td>-203 to -88</td>
</tr>
<tr>
<td>RUNX3 MSP-U</td>
<td>TTACCGCGGTTGTTTTCGCGTTT</td>
<td>AAAAAATCACGTACTTTTATTAACCAATCG</td>
<td>113</td>
<td>53</td>
<td>-203 to -88</td>
</tr>
<tr>
<td>SCGB3A1 MSP-M</td>
<td>TCAGCGGTTGTTTTCGCGTTT</td>
<td>AAATCCGTCGCTACCGCGC</td>
<td>135</td>
<td>61</td>
<td>-55 to +119</td>
</tr>
<tr>
<td>SCGB3A1 MSP-U</td>
<td>TCAGCGGTTGTTTTCGCGTTT</td>
<td>AAATCCGTCGCTACCGCGC</td>
<td>135</td>
<td>61</td>
<td>-55 to +119</td>
</tr>
<tr>
<td>SCGB3A1 BS</td>
<td>GTAGTAGGAAGTTGGTTAGGGTA</td>
<td>AAAAAATCACGTACTTTTATTAACCAATCG</td>
<td>206</td>
<td>54</td>
<td>-142 to +64</td>
</tr>
<tr>
<td>SMARCC2 MSP-M</td>
<td>TTAGCGGTTGTTTTCGCGTTT</td>
<td>CCAGCGGACGCTGCTACGCTACAG</td>
<td>157</td>
<td>66</td>
<td>-127 to +30</td>
</tr>
<tr>
<td>SMARCC2 MSP-U</td>
<td>TTAGCGGTTGTTTTCGCGTTT</td>
<td>CCAGCGGACGCTGCTACGCTACAG</td>
<td>157</td>
<td>66</td>
<td>-127 to +30</td>
</tr>
<tr>
<td>SORBS1 MSP-M</td>
<td>GTCGGGCGTTGTTTTCGCGTTT</td>
<td>CTAGCGGACGCTGCTACGCTACAG</td>
<td>165</td>
<td>68</td>
<td>-130 to +35</td>
</tr>
<tr>
<td>SORBS1 MSP-U</td>
<td>GTCGGGCGTTGTTTTCGCGTTT</td>
<td>CTAGCGGACGCTGCTACGCTACAG</td>
<td>165</td>
<td>68</td>
<td>-130 to +35</td>
</tr>
</tbody>
</table>

MSX1, MSX2, RASSF1A, RUNX3, and SCGB3A1 by MSP, a method allowing for distinction between unmethylated and methylated alleles of the studied genes (Herman et al., 1996). After bisulphite treatment of DNA, which converts unmethylated but not methylated cytosines to uracil, DNA was amplified by PCR using primers specific to methylated and unmethylated sequences (Table 1). MSP was in addition performed for the new candidate genes for methylation, identified by cDNA microarrays; CGG triplet repeat binding protein 1 (CGGBP1), cell growth regulator with ring finger domain 1 (CGRRF1), SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2 (SMARCC2), sorbin and SH3 domain containing 1 (SORBS1), and xeroderma pigmentosum, complementation group A (XPA). Human placental DNA (Sigma), which we treated in vitro with SssI methyltransferase (New England Biolabs Inc., Beverly, MA), served as positive control for methylated templates, whereas bisulphite modified DNA from normal lymphocytes was used as negative control.
The methylation status of HOXA9 and SCGB3A1 were additionally analyzed by bisulphite genomic sequencing, as previously described (Fraga et al., 2005). Six clones were sequenced from each sample to establish the methylation status of individual CpG dinucleotides. See Table 1 for primer sequences.

For the bisulphite treatment (Grunau et al., 2001; Fraga et al., 2002), 1.3 μg DNA in a total volume of 50 μL H2O was the starting material. To denature the DNA, 3.0 M NaOH (MERC, Damstadt, Germany) was added to a final concentration of 0.3 M and the sample was incubated at 37°C for 15 minutes. Thirty-three μL 20 mM Hydroquinone (Sigma) was added to the sample before sodium bisulphite (Sigma) pH 5.0 was mixed in to a final concentration of 3.7 M. The samples were incubated under aluminum foil for 16-17 hours at 50°C. After using the WIZARD DNA Clean-up kit (Promega Ltd., Southampton, UK), the DNA was eluted in 100 μL water. To complete the nucleotide conversion of unmethylated C to U, 3.0 M NaOH was added to a final concentration of 0.3 M before the sample was incubated 15 minutes at 37°C. To precipitate the bisulphite treated DNA, 1 µL Glycogeno (10 mg/ml; Ambion Ltd., Huntingdon, Cambridgeshire, UK) was added to each sample together with NH4Ac to a final concentration of 0.33 M. Finally, 900 µL 100% ethanol was mixed in and the sample was incubated at -80°C over night. The DNA was recollected by centrifugation, and washed in 70% ethanol before dried and dissolved in 30 μL water.

cDNA microarrays
Total RNA was isolated from the cell lines TERA2 and NCCIT and their 5-aza-2’-deoxycytidine treated counterparts using the RNAeasy mini kit (Qiaogen). The RNA was DNase treated according to standard procedures (Qiaegen), and eluted in 0.1% DEPC water. To evaluate the RNA quality, the RNA was separated by formaldehyde 1.2% agarose gel electrophoresis according to the RNeasy mini protocol (Qiaegen). The RNA concentration was determined using a Gene Quant Pro spectrophotometer (Biochrom Ltd. Cambridge, England). Labeled cDNA was synthesized from 30 μg RNA in a hexamer-primed polymerization with SupertScriptIII® RNase H reverse transcriptase (Invitrogen, Carlsbad, CA) in the presence of either Cy3 (test) or Cy5 (reference) labeled dUTP (Amersham Pharmacia, Piscataway, NJ). The Cy3-labeled test cDNA from the various cell lines and Cy5-labeled reference cDNA were mixed and simultaneously hybridized to the cDNA microarray. The hybridizations were done on in-house made microarrays (CNIO, Madrid, Spain) containing 13056 spots representing 7691 unique cDNA clones from 3686 genes. The human universal reference RNA (Stratagene, CA, USA) was used as a common reference for all samples. The fluorescence intensities of the spots were detected by a laser confocal scanner (Agilent Technologies, Palo Alto, CA). For each array element, a ratio between the background subtracted relative fluorescence intensities of the test and reference was calculated. The ratios were normalized using GenePix Pro (6.0; Axon Instruments, Union City, Ca, USA). Array elements upregulated two or more times after 5-aza-2’-deoxycytidine treatment were considered to be potential targets for DNA methylation.

Statistics
All 2 x 2 contingency tables were analyzed using Fisher’s exact test. The Mann –Whitney test was performed when appropriate. All P-values were derived from two-tailed statistical tests using the SPSS 11.5 software.

Acknowledgements
The study was funded by a grant (A95068, RAL) from the Norwegian Cancer Society supporting GEL as a Research Fellow and RIS as a Post-Doctoral Fellow. This study was also supported by the Rakel and Otto Kr. Bruun’s legacy, University of Oslo, Norway. The authors are grateful to Dr. Gunnar Brunborg at the Norwegian Institute of Public Health, Oslo, Norway for providing the normal testis DNA from organ donors.

References


Paper IV

Guro E Lind, Lin Thorstensen, Tone Løvig, Gunn I Meling, Richard Hamelin, Torleiv O Rognum, Manel Esteller, and Ragnhild A Lothe

A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines

A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines

Guro E Lind¹, Lin Thorstensen¹, Tone Løvig², Gunn I Meling³, Richard Hamelin⁴, Torleiv O Rognum², Manel Esteller⁵ and Ragnhild A Lothe*¹

Address: ¹Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital, 0310 Oslo, Norway, ²Institute of Forensic Medicine, The National Hospital, University of Oslo, Norway, ³The University Hospital of Akershus, Akerhus, Norway, ⁴INSERM, Paris, France and ⁵Cancer Epigenetics Laboratory, the Spanish National Cancer Centre (CNIO), Madrid, Spain

Email: Guro E Lind - guroli@ulrik.uio.no; Lin Thorstensen - l.c.thorstensen@medisin.uio.no; Tone Løvig - tlovig@labmed.uio.no; Gunn I Meling - gi@meling.net; Richard Hamelin - richard.hamelin@cephb.fr; Torleiv O Rognum - t.o.rognun@labmed.uio.no; Manel Esteller - mesteller@cnio.es; Ragnhild A Lothe* - rlothe@radium.uio.no

* Corresponding author

Abstract

Background: Tumor cell lines are commonly used as experimental tools in cancer research, but their relevance for the in vivo situation is debated. In a series of 11 microsatellite stable (MSS) and 9 microsatellite unstable (MSI) colon cancer cell lines and primary colon carcinomas (25 MSS and 28 MSI) with known ploidy stem line and APC, KRAS, and TP53 mutation status, we analyzed the promoter methylation of the following genes: hMLH1, MGMT, p16INK4a (CDKN2A α-transcript), p14ARF (CDKN2A β-transcript), APC, and E-cadherin (CDH1). We compared the DNA methylation profiles of the cell lines with those of the primary tumors. Finally, we examined if the epigenetic changes were associated with known genetic markers and/or clinicopathological variables.

Results: The cell lines and primary tumors generally showed similar overall distribution and frequencies of gene methylation. Among the cell lines, 15%, 50%, 75%, 65%, 20% and 15% showed promoter methylation for hMLH1, MGMT, p16INK4a, p14ARF, APC, and E-cadherin, respectively, whereas 21%, 40%, 32%, 38%, 32%, and 40% of the primary tumors were methylated for the same genes. hMLH1 and p14ARF were significantly more often methylated in MSI than in MSS primary tumors, whereas the remaining four genes showed similar methylation frequencies in the two groups. Methylation of p14ARF, which indirectly inactivates TP53, was seen more frequently in tumors with normal TP53 than in mutated samples, but the difference was not statistically significant. Methylation of p14ARF and p16INK4a was often present in the same primary tumors, but association to diploidy, MSI, right-sided location and female gender was only significant for p14ARF. E-cadherin was methylated in 14/34 tumors with altered APC further stimulating WNT signaling.

Conclusions: The present study shows that colon cancer cell lines are in general relevant in vitro models, comparable with the in vivo situation, as the cell lines display many of the same molecular alterations as do the primary carcinomas. The combined pattern of epigenetic and genetic aberrations in the primary carcinomas reveals associations between them as well as to clinicopathological variables, and may aid in the future molecular assisted classification of clinically distinct stages.
Background
During the last decade, epigenetic changes have been reported in many cancers and they are now recognized to be at least as common as genetic changes [1]. Aberrant methylation of cytosine located within the dinucleotide CpG is by far the best-categorized epigenetic change. The genome of the cancer cell demonstrates global hypomethylation [2,3] as well as regional promoter hypermethylation of several tumor suppressor genes [4]. Hypermethylation of selected CpG sites within CpG islands in the promoter region of genes is associated with loss of gene expression and is observed in both physiological conditions, such as X chromosome inactivation [5], and neoplasia [6]. By inactivating various tumor suppressor genes, this epigenetic modification can affect many important cellular processes, such as the cell cycle (RB, p15INK4b, p16INK4a), the TP53 pathway (p14ARF), the WNT signaling pathway (APC, E-cadherin), DNA repair (MGMT, hMLH1, BRCA1), apoptosis (DAPK), and the metastasizing process (E-cadherin, TIMP3) (reviewed in [1,7,8]).

Development of colorectal cancer through various morphological stages has been linked to several genetic and epigenetic changes. The majority of carcinomas have several chromosomal aberrations, a phenotype often referred to as chromosomal instability. Approximately 15% of the tumors are near diploid but exhibit microsatellite instability (MSI), seen as genome-wide short nucleotide insertions and deletions [9]. This phenotype is caused by a defect DNA mismatch repair system [9]. Subgroups of both types of colorectal carcinomas reveal aberrant methylation of tumor suppressor genes leading to lack of expression [10,11].

Human cancer cell lines are important tools in cancer research. Their commercial availability and unrestrained growth make them well suited for in vitro studies. Although many of the known genetic aberrations in colon cancer cell lines have been comprehensively described [12], several of these cell lines have not been analyzed for methylation status of pathogenetically important target genes.

The frequencies of both methylation and gene mutation differ among various studies of cell lines and primary tumors. The genome characteristics, profiles of gene mutations, and methylation status are rarely reported in the same samples, let alone in large series. In the present report we address these potentially connected pathogenetic mechanisms by presenting methylation profiles of a set of genes in a series of MSI and microsatellite stable (MSS) colon cancer cell lines and primary colorectal carcinomas. The methylation profiles are compared with various known genetic and clinicopathological features of the same series.

Results
Methylation status of target genes in colon cancer cell lines
The colon cancer cell line methylation-specific PCR (MSP) results are summarized in Table 1 and Figure 1a. Among the MSI cell lines 3/9, 5/9, 7/9, 8/9, 2/9, and 2/9 showed promoter hypermethylation of hMLH1, MGMT, p16INK4a, p14ARF, APC, and E-cadherin, respectively, whereas 0/11, 5/11, 8/11, 5/11, 2/11, and 1/11 of the MSS cell lines were hypermethylated for the same genes (Table 2). Hence, the cell lines with MSI generally showed higher methylation frequencies than did the MSS cell lines (Figures 1a, 2a). In most cases, methylation of the target genes was biallelic, but in 10 of the 20 cell lines, monoallelic methylation (detection of both methylated and unmethylated MSP gel bands) was found for one or more of the genes (Table 1). The MSS V9P was the only cell line unmethylated for all six genes analyzed.

Methylation status of target genes in primary colorectal carcinomas. Comparison with colon cancer cell lines
Methylation status was assessable in more than 99% of the total number of analyses (53 tumors × 6 genes = 318 analyses).

The results of the methylation analyses of 53 primary colorectal carcinomas (25 MSS and 28 MSI) are shown in Table 2 and illustrated in Figures 1b and 2b. All the methylated primary tumors examined showed an unmethylated band in addition to the methylated one, probably due to the presence of normal cells. The methylation frequencies varied from 0% among MSS tumors at the hMLH1 promoter to 61% among the MSI tumors for the p14ARF gene (Table 2).

Several of the primary tumor samples displayed widespread CpG island methylation (Figure 1b). Eighteen of 52 tumors (35%) were methylated in 3 or more of the 6 genes analyzed. Only 5/52 (10%) of the tumor samples did not show hypermethylation in any of the genes analyzed. We saw no statistical difference in the number of methylated target genes in colon cancer cell lines versus colorectal primary tumors (Mean Rank 32 for primary tumors versus 38 for cell lines, P = 0.231, Mann-Whitney test).

Methylation profiles compared with genetic characteristics
The methylation status of the primary tumors was compared with genetic characteristics of the same tumors (Table 3). In general, higher frequencies of gene methylation were found among diploid than among aneuploid tumors, reflecting the MSI status, but the differences reached statistical significance only for p14ARF (P < 0.001) and hMLH1 (P = 0.015). Sixteen of 49 primary tumors
harbored TP53 mutations, and all of the tumors with TP53 mutations also harbored unmethylated hMLH1 ($P = 0.009$). p14$^{ARF}$ hypermethylation was less common in tumors with wild type TP53 than in tumors with mutated TP53, although this was not statistically significant ($P = 0.127$). Four tumors displayed a G:C to A:T TP53 mutation and three of them simultaneously harbored a methylated MGMT gene. Overall, the presence of KRAS mutations was not associated with the methylation status of the genes analyzed. Among the 20 tumors with p14$^{ARF}$ methylation, 10 were also methylated at the adjacent p16$^{INK4a}$ gene ($P = 0.067$). Finally, the APC promoter was methylated in 17/53 (32%) tumors, and 8/17 (47%) tumors displayed both APC mutation and methylation.

Among the tumors with widespread methylation (3 or more methylated genes), 13/18 (72%) tumors demonstrated MSI, whereas 5/24 (21%) were MSS ($P = 0.080$). We found no statistically significant associations between

Table 1: Promoter methylation of colon cancer cell lines. MSI, microsatellite instable; MSS, microsatellite stable; U, unmethylated; M, methylated. The references give results in agreement with our own data except when the reference is underlined. Note that reference 15 does not use the category monoallelic methylation, but reports the promoters only as methylated or unmethylated.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>hMLH1</th>
<th>MGMT</th>
<th>p16$^{INK4a}$</th>
<th>p14$^{ARF}$</th>
<th>APC</th>
<th>E-Cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Col15</td>
<td>M$^{12}$</td>
<td>M</td>
<td>M$^{12}$</td>
<td>M</td>
<td>U/M</td>
<td>U</td>
</tr>
<tr>
<td>HCT15</td>
<td>U$^{12,13,14,15}$</td>
<td>U/M$^{16}$</td>
<td>M$^{12,14,15}$</td>
<td>M$^{14,15,17}$</td>
<td>U</td>
<td>U$^{15}$</td>
</tr>
<tr>
<td>HCT116</td>
<td>U$^{12,13,14,15,18,20,21,22}$</td>
<td>U/M$^{15,20}$</td>
<td>U/M$^{12,14,20,21,22,23}$</td>
<td>U/M$^{13,21,24,25}$</td>
<td>U/M</td>
<td>U</td>
</tr>
<tr>
<td>LoVo</td>
<td>U$^{12,13,14,15,18,22,26}$</td>
<td>U/M$^{15,31}$</td>
<td>M$^{12,14,15,22}$</td>
<td>M$^{14,15,24,25}$</td>
<td>U</td>
<td>U$^{15}$</td>
</tr>
<tr>
<td>LS174T</td>
<td>U$^{12,13,18,22}$</td>
<td>U/M</td>
<td>U$^{12,22}$</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>RKO</td>
<td>M$^{15,18,19,20,22,26}$</td>
<td>U$^{15,20}$</td>
<td>M$^{15,20,22,27}$</td>
<td>M$^{15,24}$</td>
<td>U</td>
<td>M$^{14}$</td>
</tr>
<tr>
<td>SW48</td>
<td>M$^{12,13,14,15,18,20,22,26,28,29}$</td>
<td>M$^{15,20,31}$</td>
<td>M$^{12,14,15,20,22,27,29}$</td>
<td>M$^{14,15,24}$</td>
<td>U</td>
<td>U$^{15}$</td>
</tr>
<tr>
<td>TC7</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U$^{12}$</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>TC71</td>
<td>U$^{12}$</td>
<td>U</td>
<td>M$^{12}$</td>
<td>U</td>
<td>U/M</td>
<td>U</td>
</tr>
<tr>
<td>MSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALA</td>
<td>U$^{12}$</td>
<td>U</td>
<td>M$^{12}$</td>
<td>U</td>
<td>M</td>
<td>U</td>
</tr>
<tr>
<td>Colo320</td>
<td>U$^{12,14,18,20}$</td>
<td>M</td>
<td>M$^{12,14,27}$</td>
<td>U$^{14}$</td>
<td>U$^{10}$</td>
<td>M</td>
</tr>
<tr>
<td>EB</td>
<td>U$^{12}$</td>
<td>U</td>
<td>M$^{12}$</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>FRI</td>
<td>U$^{12}$</td>
<td>U/M</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U/M</td>
<td>U</td>
</tr>
<tr>
<td>HT29</td>
<td>U$^{12,13,14,15,18,22,26,30}$</td>
<td>U$^{15,31,32,33}$</td>
<td>M$^{12,14,15,21,22,27}$</td>
<td>U$^{14,15,21,24}$</td>
<td>U$^{10}$</td>
<td>U$^{15}$</td>
</tr>
<tr>
<td>IS1</td>
<td>U$^{12,21}$</td>
<td>U</td>
<td>M$^{12,21}$</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>IS2</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U/M$^{12}$</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>IS3</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U$^{12}$</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>LS1034</td>
<td>U$^{12,12}$</td>
<td>U/M</td>
<td>U/M$^{12}$</td>
<td>M</td>
<td>U/M</td>
<td>U</td>
</tr>
<tr>
<td>SW480</td>
<td>U$^{12,13,14,15,19,21,22,26,30}$</td>
<td>U/M$^{15,15}$</td>
<td>M$^{12,14,15,21,22,27}$</td>
<td>U$^{14,15,21,24,25}$</td>
<td>U$^{10}$</td>
<td>U$^{15}$</td>
</tr>
<tr>
<td>V9P</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U$^{12}$</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>

Table 2: Methylation frequencies among MSS and MSI colon cancer cell lines and primary colorectal tumors. Abbreviations; MSS, microsatellite stable; MSI, microsatellite instable; CRC, colorectal cancer; U, unmethylated; M, methylated. Note that the calculated methylation frequencies of the MSS cell lines includes results from three cell lines derived from the same patient.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cell lines</th>
<th>CRCs</th>
<th>Cell lines</th>
<th>CRCs</th>
<th>Cell lines</th>
<th>CRCs</th>
<th>Cell lines</th>
<th>CRCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMLH1</td>
<td>0/11 (0%)</td>
<td>0/25 (0%)</td>
<td>3/9 (33%)</td>
<td>11/28 (39%)</td>
<td>3/20 (15%)</td>
<td>11/53 (21%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGMT</td>
<td>5/11 (45%)</td>
<td>10/25 (40%)</td>
<td>5/5 (56%)</td>
<td>11/28 (39%)</td>
<td>10/20 (50%)</td>
<td>21/53 (40%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p16$^{INK4a}$</td>
<td>8/11 (73%)</td>
<td>7/25 (28%)</td>
<td>7/7 (100%)</td>
<td>10/28 (36%)</td>
<td>15/20 (75%)</td>
<td>17/53 (32%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p14$^{ARF}$</td>
<td>5/11 (45%)</td>
<td>3/24 (12%)</td>
<td>8/9 (89%)</td>
<td>17/28 (61%)</td>
<td>13/20 (65%)</td>
<td>20/52 (38%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>2/11 (18%)</td>
<td>7/25 (28%)</td>
<td>2/9 (22%)</td>
<td>10/28 (36%)</td>
<td>4/20 (20%)</td>
<td>17/53 (32%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-cadherin</td>
<td>1/11 (9%)</td>
<td>10/24 (42%)</td>
<td>2/9 (22%)</td>
<td>11/28 (39%)</td>
<td>3/20 (15%)</td>
<td>21/53 (40%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

http://www.molecular-cancer.com/content/3/1/28
tumors with widespread methylation and presence of TP53, KRAS, or APC mutations.

**Methylation profiles and clinicopathological features**

The clinicopathological features and methylation status of the primary tumors are summarized in Table 3. We saw more methylation among tumors from females than in those from males for both hMLH1 (P = 0.043) and p14ARF (P = 0.050). Tumors from patients younger than the mean age (68 years) had a lower methylation frequency for p16INK4a than did tumors from older patients, although this was not statistically significant (P = 0.074). There was a strong association between methylation and right-sided tumor location as 10/11 (91%) tumors methylated in hMLH1 and 12/19 (63%) of the tumors methylated in p14ARF were located in the right side of the colon (P < 0.001 and P = 0.005, respectively). There was no statistically significant association between methylation and histological grade. Most of the tumors with APC methylation (13/17, 76%) belonged to the Dukes’ B group, but the differences were not statistically significant (P = 0.068).

Tumors with widespread methylation (≥ 3 loci) are associated with right-sided localization; 10/17 (59%), versus 5/17 (29%) left-sided (P = 0.035). We saw no statistically significant associations between presence of widespread methylation and the remaining clinicopathological variables included in the present study.

**Discussion**

Tumor cell lines are commonly used as experimental tools in cancer research, including studies designed to assess epigenetic changes. But whereas the genetic aberrations of colon cancer cell lines have been comprehensively described [12], the methylation profiles of potential target genes in the same or similar cell lines are often described only sparingly. A literature survey of the 20 colon cancer cell lines and their methylation status analyzed in this study showed that some cell lines and genes had been extensively studied, whereas others were left undescribed (Table 1). For half of the cell lines included in the present study, both methylated and unmethylated alleles have been found for one or more of the genes studied. As non-neoplastic cells are not found in cultured cancer cell lines, this can not be caused by the presence of normal cells, and although several biological and technical explanations may exist, allele specific methylation seems the most likely interpretation [23,34]. In contrast, admixture of normal cells, tumor heterogeneity and/or monoallelic methylation may explain the coexistence of unmethylated and methylated bands in primary tumors.
It has been debated for some time whether cell lines are more frequently methylated than primary tumors [35]. Regarding overall CpG island hypermethylation, cancer cell lines have in general demonstrated an increased frequency of hypermethylation compared with primary tumors [15]. However, only a limited number of the genes analyzed have shown a statistically significant difference in methylation frequency [15]. Among several cancer types examined, colon cancer cell lines have been shown to resemble the most their respective primary tumor in this respect [36]. For the cell lines and primary tumors included in this study, the fraction of MSI and MSS samples was about the same and we saw no statistical difference in the overall number of methylated target genes in colon cancer cell lines versus colorectal primary tumors. Seemingly, large methylation percentage differences for individual genes were seen (Table 2) but they were statistically significant only for $p16^{\text{INK4a}}$ methylation, independent of MSI stratification. Comparisons of in vitro tumor cells with primary tumors of each subtype (MSS and MSI) have also shown similar frequencies of TP53, KRAS and APC mutations [12] and ploidy stem line [37], which further supports the conclusion that the in vitro system is a suitable experimental tool that closely reflect the in vivo situation.

Previously reported variations in promoter hypermethylation frequencies of different tumor suppressor genes in colon cancer cell lines and colorectal primary tumors.

![Figure 2](image_url)

**Figure 2**

*Promoter hypermethylation in colon cancer cell lines and colorectal primary tumors.* Methylation was evaluated by methylation-specific PCR (MSP). A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. The upper panel (a) illustrates the methylation status of all the loci analyzed in a MSI cell line (RKO) and a MSS cell line (HT29). The lower panel (b) shows the methylation status of representative primary colorectal tumors. Abbreviations: NB, normal blood (positive control for unmethylated samples); MP, methylated placenta (positive control for methylated samples); neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product.
<table>
<thead>
<tr>
<th></th>
<th>hMLH1</th>
<th>MGMT</th>
<th>p16^{INK4a}</th>
<th>p14^{ARF}</th>
<th>APC</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>U</td>
<td>M</td>
<td>U</td>
<td>M</td>
<td>U</td>
</tr>
<tr>
<td>Gen. Characteristics</td>
<td>Ploidy</td>
<td>Diploid</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Aneuploid</td>
<td>1</td>
<td>22</td>
<td>11</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.02</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.001</td>
<td>NS</td>
</tr>
<tr>
<td>MSI-status</td>
<td>Wild type</td>
<td>11</td>
<td>17</td>
<td>11</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Mutation</td>
<td>0</td>
<td>16</td>
<td>7</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.01</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>wt-t non G-A mutation</td>
<td>11</td>
<td>33</td>
<td>15</td>
<td>29</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>G-A mutation</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>K-Ras</td>
<td>Wild type</td>
<td>8</td>
<td>19</td>
<td>13</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Mutation</td>
<td>1</td>
<td>14</td>
<td>6</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>0.08</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>wt-t non G-A mutation</td>
<td>8</td>
<td>23</td>
<td>15</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>G-A mutation</td>
<td>1</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>APC</td>
<td>Wild type</td>
<td>7</td>
<td>19</td>
<td>12</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Mutation</td>
<td>3</td>
<td>23</td>
<td>8</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Clin. and Path. Features</td>
<td>Sex</td>
<td>Male</td>
<td>2</td>
<td>23</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>9</td>
<td>19</td>
<td>12</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.04</td>
<td>NS</td>
<td>NS</td>
<td>0.05</td>
<td>NS</td>
</tr>
<tr>
<td>Age (years)</td>
<td>&lt;68</td>
<td>2</td>
<td>21</td>
<td>10</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>≥68</td>
<td>9</td>
<td>21</td>
<td>11</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.09</td>
<td>NS</td>
<td>0.07</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Location</td>
<td>Right</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>1</td>
<td>19</td>
<td>8</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Rectum</td>
<td>0</td>
<td>14</td>
<td>6</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>&lt;0.001*</td>
<td>NS</td>
<td>0.05</td>
<td>NS</td>
<td>0.01</td>
</tr>
<tr>
<td>Histologic grade</td>
<td>Poorly differentiated</td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Moderately differentiated</td>
<td>7</td>
<td>30</td>
<td>13</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Well differentiated</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Dukes' classification</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5</td>
<td>22</td>
<td>10</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>0.07</td>
</tr>
</tbody>
</table>
colorectal cancer can be explained by various ratios of MSI versus MSS samples in the series analyzed, different methods for analyzing methylation, the inter-individual variation in scoring of methylated samples, incomplete bisulphite modification, tumor heterogeneity, and the fact that different parts of the gene promoter region in question have been analyzed. In the present study, we used primer sets known to only detect methylation in tumor cells, never in normal tissues from the same patients [24,31,38-42]. The promoter hypermethylation in these areas has also shown an impressive correlation with lack of protein expression, confirming that these are essential regions for gene expression [24,31,38-42]. The hMLH1 primers we designed amplify a region of the promoter, in which methylation invariably correlates with the lack of hMLH1 expression [18,43,44]. Methylation of this region has only been detected in tumor cells and not in normal mucosa [18,43,44].

As expected, the MSI primary tumors showed more methylation overall than did the MSS group. However, this was only significant for the hMLH1 and p14ARF genes, whereas the four additional genes analyzed revealed similar methylation frequencies in the MSS and MSI groups. Promoter methylation of the hMLH1 gene was, not surprisingly, found only in tumors and cell lines with MSI, not in the MSS samples. The MSS tumors and cell lines per definition contain functional hMLH1 protein, and transcriptional silencing of hMLH1 by hypermethylation is known to be the main cause of MSI in sporadic CRC [26,28,45]. Also p14ARF methylation may have a specific role in MSI tumors, since it seems to be most often inactivated in tumors with wild type TP53 (see below). However, the relatively high methylation frequencies of the remaining analyzed genes, and also their overall similar frequency in MSI and MSS samples, imply that they are important in colorectal carcinogenesis independently of tumor site and MSI status.

Inactivation of tumor suppressor genes by promoter hypermethylation has been recognized to be at least as common as gene disruption by mutation in tumorigenesis [1]. Indeed, most types of primary tumors harbor several genes inactivated in this way and some genes, like p16INK4a, have been reported to be methylated consistently in most tumor types analyzed [46]. In colorectal carcinomas, the reported p16INK4a methylation frequencies vary from 18% [47] to 50% [48] with most of the observations centered around 36–40% [11,27,46,49-51], i.e., slightly higher than our result. Both p16INK4a and p14ARF are more commonly methylated in tumors with MSI than in MSS [10,11,51-53], although we found that the methylation frequency of p14ARF is higher than that for p16INK4a in MSI colorectal carcinomas.

The DNA repair protein MGMT is able to remove pro-mutagenic alkyl groups from O6-guanine by an irreversible transfer to an internal cysteine residue [54]. Left unrepaired, the alkylated O6-guanine has a tendency to base pair with thymine during replication, thereby introducing a G:C to A:T transition mutation in the DNA [55]. Inactivating promoter hypermethylation of the MGMT gene has previously been reported to be associated with G:C to A:T mutations in the tumor suppressor gene TP53 [56] and the proto-oncogene KRAS [57,58]. Our data support this assumption for TP53 but seemingly not for KRAS, although no certain conclusions can be drawn from the limited number of samples with G:C to A:T mutations.

The p14ARF protein interacts in vivo with the MDM2 protein, neutralizing MDM2’s inhibition of TP53 [59]. Less hypermethylation of p14ARF in tumors with mutated TP53 than in tumors with wild type TP53 has been reported previously [24]. Additionally, several reports have described an inverse relationship between MSI and TP53 mutation in colorectal carcinomas [60-62]. The frequent methylation we report for the p14ARF gene in MSI tumors with few TP53 mutations is in agreement with a recent study [53] and supports the existence of this alternative pathway for TP53 inactivation.

Inactivation of the APC gene is frequent in colorectal and other gastrointestinal carcinomas, usually by truncating mutations [63,64]. An alternative mechanism to inactivate the gene in colorectal tumors is by promoter methylation, and we report a frequency of APC methylation in the upper range of what has been seen in previous studies [51,65,66]. Somatic mutations in APC are common in colorectal cancer [67,68] and, similar to what has been seen by others [12,22,69], almost half of the tumors displaying APC mutations in our study were also methylated. We have not looked at allele-specific mutation, but methylation and mutation in the same tumor might reflect one mutated allele and methylation of the other, in accordance with Knudson’s two hit hypothesis. This has previously been demonstrated for APC in colorectal cancer samples by Esteller et al [65]. APC has a central role in the WNT signaling pathway, which is suggested to play a part in colorectal carcinogenesis by its constitutive activation. Activation of this pathway results in increased transcription levels of genes like MYC and CCND1 (cyclin D1) further stimulating cell proliferation [63]. Among the 52 successfully analyzed primary tumors in this study, 35 had altered APC caused by methylation (n = 17) and/or gene mutation (n = 26). The E-cadherin gene was also methylated in 14/34 tumors with altered APC, presumably further stimulating WNT signaling [63]. Interestingly, APC methylation seemed to be more common in Dukes B stage tumors.
The present study confirms that methylation of hMLH1 in sporadic carcinomas is associated with proximal tumor location in the large bowel [14,21,45,70], as above 90% of the primary tumors harboring a methylated hMLH1 promoter were taken from the right side of the colon. An association between sporadic proximal colon carcinomas and methylation has also been reported for p16INK4a and p14ARF [14,21,45]. Among our 53 primary tumors, we can only confirm this statistically for p14ARF. However, p16INK4a demonstrated the same tendency. Both hMLH1 and p14ARF are strongly associated with MSI and MSI is in turn strongly associated with proximal tumor location [71,72], hence, it is not unexpected that the methylation of both genes is associated with proximal location.

When it comes to gene methylation and its association with other clinicopathological features, contradictory results have been reported. Our observation that methylation of p14ARF does not exclude p16INK4a methylation, is in accordance with previous studies [21,24]. Correlation of p16INK4a or p14ARF methylation with female gender and increased age has been described in some studies [14,47] but not in others [11,21,24]. We found such an association between female gender and methylation of p14ARF and hMLH1, but not of p16INK4a. We also found a weak association between p16INK4a methylation and increasing age. This potential age-specific methylation was not confirmed for any of the other genes studied. The gender-associated methylation of hMLH1 has previously been described [73,74] and might explain the increased prevalence of colorectal tumors of the MSI type in the female patient group [74].

Like Toyota et. al [51], we found no statistically significant associations between tumors with widespread methylation and age, gender, or stage of the colorectal cancer.

Conclusions
The data presented here demonstrate that multiple genes are methylated in colorectal carcinomas. This underlines the important role epigenetic inactivation of tumor suppressor genes plays during the process of tumor development. Epigenetic changes in colon cancer cell lines are overall comparable with those of primary carcinomas of the large bowel, which make the cell lines relevant models for the in vivo situation. The methylation profile of specific genes, in particular hMLH1 and p14ARF, has strong associations with genetic and clinicopathological features and might be related to biologically distinct subsets of colorectal tumors.

Methods

Patients and cell lines
Fifty-three primary colorectal carcinomas from 52 patients, including 25 MSS tumors and 28 MSI tumors, were submitted to methylation analyses. One of the tumors was from a patient with hereditary non-polyposis colorectal cancer (HNPCC), whereas the rest of the cases were sporadic [71]. The tumors have known DNA ploidy pattern [75], MSI status [76], as well as mutation status for TP53, KRAS and APC [62,64,77]. The genetic and clinicopathological variables are found in Table 3. Twenty colon cancer cell lines, 11 MSS and 9 MSI, were also included in the study. These cell lines have previously been characterized for MSI status [61,78-80], 31 different genetic alterations [12], and total genome profiles by Kleivi et. al [37]. The primary tumors included in the present study are from a series of carcinomas evaluated to contain a mean number of 84% tumor cells [81]. The DNA was extracted by standard phenol -chloroform procedure.

Methylation-specific PCR (MSP)
Promoter methylation was studied in hMLH1, MGMT, p16INK4a, p14ARF, APC and E-cadherin by MSP, a method that distinguishes unmethylated from methylated alleles of a given gene [38]. After bisulphite treatment of DNA, which converts unmethylated but not methylated cytosines to uracil, DNA is amplified by PCR using primers specific to methylated and unmethylated sequences.

One or two µg DNA from each sample was modified as described [82]. Previously reported primer sets were used for amplification of the MGMT [31,82], p16INK4a [38,82], p14ARF [24], APC [39,40] and E-cadherin fragments [41](island 3). The primers for amplifying unmethylated and methylated hMLH1 fragments were designed in accordance with hMLH1 promoter methylation and gene expression studies [18,44]. All primer sets (see Additional file 1) were purchased from Medprobe AS (Oslo, Norway).

All the PCRs were carried out in a total volume of 25 µl containing 1 × PCR Buffer (15mM MgCl₂ or no MgCl₂; QIAGEN Inc., Valencia, CA), 200 µM dNTP (Amersham Pharmacia Biotech Products Inc., Piscataway, NJ), and 0.625 U HotStarTaq DNA Polymerase (QIAGEN). PCR products were loaded onto 7.5% polyacrylamide gels, stained with ethidium bromide, and visualized by UV illumination. An independent second "methylated reaction" of the MSP was performed for all the samples included in the present study. In cases with diverging results from the two rounds of MSP, we did a third independent MSP round.

Human placental DNA (Sigma Chemical Co., St. Louis, MO) treated in vitro with SssI methyltransferase (New England Biolabs Inc., Beverly, MA) was used as a positive control for MSP of methylated alleles, whereas DNA from normal lymphocytes was used as a control for unmethylated alleles. Water was used as a negative PCR control in both reactions.
Authors' contributions
GEL cultured and isolated DNA from all cell lines and carried out the MSP analyses of these and of the patient samples. GEL interpreted the results, performed the statistics and drafted the manuscript. LT participated in the study design, scored the MSP results independently of author 1, and contributed to manuscript preparation. TL was responsible for the update of the APC mutation status in the cohort. GIM and TOR have collected the series of patient samples. GEL interpreted the results, performed the statistics and contributed to manuscript preparation. TL was responsible for its design and coordination, and contributed in the evaluation of the results and in preparation of the manuscript. All authors have read and approved of the final manuscript.

Additional material

Acknowledgements
We thank Professor Sverre Heim for critically reading the manuscript. GEL is a Research Fellow and LT a Post-Doctoral Fellow of the Norwegian Cancer Society. TL is Post-Doctoral Fellow of the Norwegian Foundation for Health and Rehabilitation. The study was funded by grants from the Norwegian Cancer Society (A95068 and D97127 RAL).

References


Identification of Novel Gene Targets in Colorectal Tumorigenesis: *ADAMTS1*, *CRABP1*, and *NR3C1*

Submitted manuscript
Identification of Novel Gene Targets in Colorectal Tumorigenesis:

**ADAMTS1, CRABP1, and NR3C1**

GURO E. LIND,* KRISTINE KLEIVI,* TRUDE H. AAGESEN,* GUNN I. MELING,§ TORKLEIV O. ROGNUM,∥ and RAGNHILD A. LOTHE*☆

*Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo, Norway, §Surgical Department, Faculty Division Akershus University Hospital, University of Oslo, Oslo, Norway,∥ Institute of Forensic Medicine, The National Hospital, University of Oslo, Norway ☆Department of Molecular Biosciences, University of Oslo, Oslo, Norway.

**Background and Aims:** Gene silencing through CpG island hypermethylation plays a vital role in embryogenesis as well as in tumorigenesis. In the present study we aimed to identify novel target genes for inactivation by promoter hypermethylation in colorectal cancer.

**Methods:** We compared the gene expression profiles of colon cancer cell lines before and after treatment with the demethylating agent 5-aza-2'-deoxycytidine using microarrays. The gene expression of responding genes was investigated in primary colorectal carcinomas and normal colon tissue by microarray analyses. Potential candidate genes were submitted to methylation specific-PCR (MSP), bisulphite sequencing, and quantitative gene expression analyses. **Results:** Sixty of the 93 array elements responding to treatment had down-regulated gene expression in carcinomas compared with normal samples. Twenty-one of these genes had CpG islands and four were analyzed for promoter methylation. Among 20 colon cancer cell lines hypermethylation of **ADAMTS1, CRABP1, NDRG1, and NR3C1** was found in 17 (85%), 18 (90%), 0 (0%), and 7 (35%), respectively. **In vitro** quantitative gene expression analyses confirmed that the transcripts of hypermethylated genes were either absent or reduced, statistically significant for two of the three genes. The methylation status of **ADAMTS1, CRABP1, and NR3C1** were also investigated in primary colorectal adenomas and carcinomas and 23/63 (37%), 7/60 (12%), and 2/63 (3%) adenomas, and 37/52 (71%), 25/51 (49%), and 13/51 (25%) carcinomas were hypermethylated for the respective genes.

**Conclusions:** **ADAMTS1, CRABP1, and NR3C1** are novel targets frequently inactivated by promoter hypermethylation in colorectal carcinomas. Their inactivation also in a subgroup of benign lesions strongly supports their importance as early events in colorectal tumorigenesis.
**Introduction**

Impaired epigenetic regulation is as common as gene mutations in human cancer (1). Both mechanisms lead to quantitative and qualitative gene expression changes causing a selective growth advantage in the cell population, which may result in a cancerous transformation. Aberrantly hypermethylated CpG islands in the gene promoter associated with transcriptional inactivation is one of the most frequent epigenetic changes in cancer.

In the large bowel the development of the adenoma- carcinoma lineage is paralleled by several genetic and epigenetic changes. One of these is the hypermethylation of $MLH1$, which leads to inactivation of the DNA mismatch repair system, and subsequent microsatellite instability (MSI) (2). Approximately 15% of all sporadic colorectal carcinomas display MSI seen as insertions and deletions of short nucleotide repeats in their near diploid genome (3). The microsatellite stable tumors (MSS) have a functional mismatch repair system, but usually display chromosomal aberrations including several common regions of loss or gain (4). Subgroups of both types of colorectal carcinomas reveal aberrant methylation of tumor suppressor genes associated with lack of expression, although the methylation frequencies are generally higher in the MSI tumor group (5). Based on the identification of two epigenetically distinct tumor groups, one in which methylation is extremely rare, and a second where the tumors harbor methylation of multiple loci, a third pathway to colorectal tumorigenesis has been suggested, the CpG island methylator phenotype (CIMP) (6). However, the original findings are supported by some reports (7; 8), but not by others (9; 10), leaving the CIMP concept highly controversial.

We have used microarray technology to compare the gene expression in colon cancer cell lines before and after treatment with 5-aza-2’-deoxycytidine, a compound that removes the methylation genome-wide. We investigated the expression of the responding genes in primary tumors and demonstrate strict selection criteria for identifying new hypermethylated target genes. Finally, we present three novel genes epigenetically inactivated in colorectal adenomas, carcinomas, and colon cancer cell lines.
Materials and Methods

Cell Cultures and Tissue Samples

Twenty colon cancer cell lines were included in the present study, nine MSI: Co115, HCT15, HCT116, LoVo, LS174T, RKO, SW48, TC7, and TC71, and 11 MSS: ALA, Colo320, EB, FRI, HT29, IS1, IS2, IS3, LS1034, SW480, and V9P (11). All cell lines were cultured in DMEF-12 medium (GIBCO, Invitrogen Carlsbad, CA) with 10% fetal bovine serum (GIBCO), 2 mM L-glutamine (GIBCO), 100 U/ml penicillin G, and 100 µg/ml streptomycin (GIBCO) and harvested before confluency. Two MSI cell lines (HCT15, SW48) and two MSS cell lines (HT29, SW480) were cultured in parallel with and without 10 µM 5-aza-2’-deoxycytidine (Sigma-Aldrich Company Ltd. Dorset, UK) in the medium for 72 hours. DNA was extracted from the cell lines and their 5-aza-2’-deoxycytidine treated counterparts by a standard phenol-chloroform procedure, and total RNA was isolated using Trizol (Invitrogen, Carlsbad, CA). The RNA quality was measured by a 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA) and the RNA concentration was determined using ND-1000 Nanodrop (NanoDrop Technologies, Wilmington, DE). DNA from 53 primary colorectal carcinomas (25 MSS and 28 MSI) from 52 patients, and three normal colon mucosa samples from individual colorectal cancer patients (derived in distance from the tumors) were additionally submitted to methylation analyses.

cDNA Microarrays

Eighteen cell lines and their 5-aza-2’-deoxycytidine treated counterparts (n = 4) were analyzed by cDNA microarrays. Labeled cDNA was synthesized from 30 µg total RNA in an oligo dT-primed polymerization with SupertScript™II RNase H reverse transcriptase (Invitrogen) in the presence of either Cy3 (test) or Cy5 (reference) labeled dUTP (Amersham Pharmacia, Piscataway, NJ). Test and reference cDNA was mixed and hybridized onto cDNA microarrays (The microarray core facilities, The Norwegian Radium Hospital, Oslo, Norway), containing 15486 spots representing 12688 unique cDNA clones from 8461 genes. The human universal reference RNA (Stratagene, CA, USA) was mixed in a 4:1 relationship with total RNA from four combined colon cancer cell lines (two MSI; HCT116 and LoVo, and two MSS; HT29 and SW480), and used as a common reference for all samples. The fluorescence intensities of the spots were detected by a laser confocal scanner (Agilent Technologies). For each array element, a ratio between the
Epigenetic target genes in colorectal cancer

background subtracted relative fluorescence intensities of the test and reference was calculated (GenePix Pro 6.0; Axon Instruments, Union City, Ca, USA). The ratios in all samples were post-processed and normalized by the lowess method using BASE (12). Array elements up-regulated two or more times after 5-aza-2'-deoxycytidine treatment in at least three of four cell lines were considered to be potential targets for DNA methylation.

Bisulphite Treatment and Methylation specific-PCR

Bisulphite treatment of DNA leads to sequence variations as unmethylated but not methylated cytosines are converted to uracil (13). DNA from cell lines and primary colorectal carcinomas was bisulphite treated as previously described (14; 15), whereas DNA from the adenomas was bisulphite treated according to the protocol of the CpGenome™DNA modification kit (Intergen Boston, MA) (16). The promoter methylation status of ADAMTS1, CRABP1, NDRG1, and NR3C1 was subsequently analyzed by methylation-specific PCR (MSP), a method allowing for distinction between unmethylated and methylated alleles (17). All primers were designed with MethPrimer (18) (with the exception of NR3C1 bisulphite sequencing primers, see acknowledgements) and their sequences are listed in Table 1, along with the product fragment length, annealing temperature, and magnesium concentration for each PCR. The fragments were amplified using the HotStarTaq DNA polymerase (QIAGEN Inc., Valencia, CA), and all analyses were confirmed with a second independent round of MSP.

Table 1. Primers used for methylation specific-PCR and bisulphite sequencing. Abbreviations: MSP, methylation specific-PCR; BS, bisulphite sequencing; M, methylated-specific primers; U, unmethylated-specific primers; Frg.size, fragment size; An.temp, annealing temperature. *The annealing temperature is given in degrees celcius.
Bisulphite Sequencing

Bisulphite sequencing allows a positive display of 5-methyl cytosines in the gene promoter after bisulphite modification as unmethylated cytosines appear as thymines, while 5-methylcytosines appear as cytosines in the final sequence (13). For ADAMTS1, CRABP1, and NR3CI we submitted a subset of the cell lines to bisulphite sequencing \((n = 13, n = 9, \text{ and } n = 10, \text{ respectively})\), based on the MSP results. The majority of the cell lines analyzed was methylated from the MSP analyses, but a minimum of one unmethylated and one monoallelic methylated cell line was also included in the panel sequenced for each gene. Primer sequences and PCR conditions are listed in Table 1. All fragments were amplified with the HotStarTaq DNA Polymerase and eluted from a 2% agarose gel by the MinElute™ Gel Extraction kit (QIAGEN). The samples were subsequently sequenced with the dGTP BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster city, CA) in an ABI Prism 377 Sequencer (Applied Biosystems). The approximate amount of methyl cytosine of each CpG site in the various fragments was calculated by comparing the peak height of the cytosine signal with the sum of the cytosine and thymine peak height signals, as previously described (19).

Quantitative Gene Expression Analyses

We used TaqMan real-time fluorescence detection (Applied Biosystems, Foster city, CA) to quantify mRNA levels in the colon cancer cell lines, as previously described (20; 21). cDNA was generated from two \(\mu\)g total RNA using the SuperScript™ II reverse transcriptase enzyme (Invitrogen) and oligo dT primers (Medprobe) according to the manufacturers’ protocol. cDNA from the genes of interest \((ADAMTS1, \text{ Hs01095534_g1}; \text{ CRABP1 Hs00171635_m1}; \text{ and } \text{NR3CI, Hs00230818_m1})\) and the endogenous controls \((ACTB, \text{ Hs99999903_m1 and } \text{PGK1, Hs00943173_gH})\) were amplified separately by the ABI Prism 7000 Sequence Detection System following the protocol recommended by Applied Biosystems. All samples were analyzed in duplicates. The quantitative expression levels were measured against a standard curve generated from dilutions of cDNA from the human universal reference RNA (Stratagene). In order to adjust for the possible various amounts of cDNA input in each PCR, we normalized the expression quantity of the target genes with the quantity of the housekeeping gene \(PGK1\),
Epigenetic target genes in colorectal cancer

Figure 1. Gene expression of potential candidates for inactivation by DNA hypermethylation in colon cancer cell lines (A), 5-aza-2’-deoxycytidine treated cell lines (B), and primary colorectal carcinomas (C). The panels illustrate the gene expression in alphabetically order as assessed by microarray analyses. Each ratio is presented by a color for visualization. The gene expression of 21 genes containing a CpG island in their promoter was induced by 5-aza-2’-deoxycytidine treatment in colon cancer cell lines (B) and was simultaneously down-regulated in primary colorectal carcinomas relative to normal colon mucosa (C). Genes submitted to methylation analyses are highlighted in red.

which was determined to be best suitable (see results).

Statistics

All 2 x 2 contingency tables were analyzed using Fisher’s exact test. For the statistical analyses of the quantitative gene expression a 2 x 3 table and Chi-square test was used. The gene expression was divided in two categories; low expression includes samples with gene expression ≤ median value for the gene in question across all cell lines; high expression includes samples with gene expression > median. Methylation status was divided in three categories: unmethylated, partial methylation, and methylated. All P values are derived from two tailed statistical tests using the SPSS 11.5 software (SPSS, Chicago, IL, USA).

Results

Identification of Novel Candidate Genes Potentially Epigenetically Deregulated in Colorectal Carcinomas

Ninety-three array elements were up-regulated two or more times after 5-aza-2’-deoxycytidine treatment in at least three out of the four cell lines analyzed (HCT15, HT29, SW48, and SW480).
Among the 93 elements, 88 were overlapping with a panel of 18 primary colorectal carcinomas and three normal colon tissue samples analyzed by oligo microarrays (Kristine Kleivi, unpublished data). The median of 60 of these genes was down-regulated across the panel of tumor samples relative to normal colon tissue and 21 of them contained a CpG island in their 5′region. Among these, we selected four candidates encoding proteins with potential roles in tumor development; ADAMTS1, CRABP1, NDRG1, and NR3C1 for promoter methylation analyses in colon cancer cell lines (Figure 1). Prior to the analyses, the microarray cDNA clones responding to 5-aza-2′-deoxycytidine treatment were sequenced, and their identity confirmed.

**Methylation Status of Novel Candidate Genes in vitro and in vivo**

The promoters of ADAMTS1, CRABP1, and NR3C1 were hypermethylated in 17/20 (85%), 18/20 (90%), and 7/20 (35%) colon cancer cell lines, respectively, whereas NDRG1 was unmethylated in all samples from the MSP analyses (Table 2 and Figure 2). Thirty-seven of 52 (71%) colorectal carcinomas were hypermethylated for ADAMTS1, with equal frequencies in MSI and MSS tumors, and 23/63 (37%) of the colorectal

![Figure 2. Representative methylation specific-PCR results from the analysis of ADAMTS1, CRABP1, NDRG1, and NR3C1 in sixteen colon cancer cell lines. A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. NB, normal blood (positive control for unmethylated samples); IVD, in vitro methylated DNA (positive control for methylated samples); neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product. Each gel panel is a merge of two to three gel panels mainly run on the same agarose gel.](image)
<table>
<thead>
<tr>
<th>Cell line</th>
<th>MSI status</th>
<th>ADAMTS1</th>
<th>CRABP1</th>
<th>NDRG1</th>
<th>NR3C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co115</td>
<td>MSI</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>HCT15</td>
<td>MSI</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U/M</td>
</tr>
<tr>
<td>HCT116</td>
<td>MSI</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>LoVo</td>
<td>MSI</td>
<td>U/M</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>LS174T</td>
<td>MSI</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>RKO</td>
<td>MSI</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>SW48</td>
<td>MSI</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>TC7</td>
<td>MSI</td>
<td>U/M</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>TC71</td>
<td>MSI</td>
<td>M</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>ALA</td>
<td>MSS</td>
<td>U</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Colo320</td>
<td>MSS</td>
<td>U</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>EB</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>FRI</td>
<td>MSS</td>
<td>U</td>
<td>U/M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>HT29</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>IS1</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>IS2</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>IS3</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>LS1034</td>
<td>MSS</td>
<td>M</td>
<td>M</td>
<td>U</td>
<td>U/M</td>
</tr>
<tr>
<td>SW480</td>
<td>MSS</td>
<td>M</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>V9P</td>
<td>MSS</td>
<td>U/M</td>
<td>M</td>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carcinomas</th>
<th>MSI and MSS*</th>
<th>ADAMTS1</th>
<th>CRABP1</th>
<th>NDRG1</th>
<th>NR3C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td>37/52 (71%)</td>
<td>25/51 (49%)</td>
<td>Nd</td>
<td>13/51 (25%)</td>
<td></td>
</tr>
<tr>
<td>MSS</td>
<td>20/28 (71%)</td>
<td>22/28 (79%)</td>
<td>Nd</td>
<td>12/28 (43%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adenomas</th>
<th>MSI and MSS+</th>
<th>ADAMTS1</th>
<th>CRABP1</th>
<th>NDRG1</th>
<th>NR3C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td>23/63 (37%)</td>
<td>7/60 (12%)</td>
<td>Nd</td>
<td>2/63 (3%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Promoter hypermethylation of candidate genes in colon cancer cell lines, primary colorectal carcinomas, and adenomas. Abbreviations: U, unmethylated; M, methylated; MSI, microsatellite stable; MSS, microsatellite unstable; Nd, not done. * MSI and MSS carcinomas grouped together. + Only two adenomas are MSI, the remaining 61 are MSS. Here they are presented together.

adenomas were hypermethylated. CRABP1 was significantly more frequently methylated among the MSI carcinomas (22/28; 79%) than the MSS group (3/23; 13%; \( P < 0.001 \)), and 7/60 (12%) of the adenomas were hypermethylated. The methylation frequency of NR3C1 was also significantly higher in MSI carcinomas (12/28; 43%) than in MSS carcinomas (1/23; 4%; \( P = 0.003 \)). Two of 63 (3%) adenomas were hypermethylated for NR3C1. Additionally, the promoters of ADAMTS1, CRABP1, NDRG1, and NR3C1 were all unmethylated in the three normal colon tissue samples.
MSP analyses of the four 5-aza-2'-deoxycytidine treated colon cancer cell lines verified that \textit{ADAMTS1}, \textit{CRABP1}, and \textit{NR3C1} were demethylated in 80% (8/10) of the originally methylated cases. \textit{ADAMTS1} was still methylated in 5-aza-2'-deoxycytidine treated SW480 cells and no change in gene expression was detected from the microarray analyses. However, in drug treated SW48 cells, \textit{NR3C1} expression was up-regulated compared with untreated cells even in the absence of visible demethylation.

**Bisulphite Sequencing Results**

Bisulphite genomic sequencing of \textit{ADAMTS1}, \textit{CRABP1}, and \textit{NR3C1} in selected cell lines showed that all cytosines at non-CpG sites were converted to thymine. This is seen in Figure 3 along with the detailed sequencing results and the initial promoter methylation status as assessed by MSP. In general, there seems to be a good association between the MSP scoring and the bisulphite sequences. Five out of nine cell lines methylated in \textit{ADAMTS1}, assessed by MSP, display only fully methylated CpG sites, whereas the remaining four cell lines display partial methylation (the presence of both methylated and unmethylated cytosine) in one to six CpG sites. Among these, the TC71 cell line has five partially methylated CpG sites including sites 17 to 19, which are fully methylated in all other cell lines methylated from MSP analyses. In the \textit{CRABP1} gene the majority of CpG sites in the methylated cell lines were also fully methylated from the bisulphite sequencing analyses. Some partial methylation was seen around CpG sites 6 to 11 as five out of the six methylated cell lines from MSP displayed both methylated and unmethylated cytosines of two to five of these CpG sites. RKO was the only methylated cell line from MSP analysis harboring partial methylation for \textit{NR3C1} and only in two of the 58 successfully amplified CpG sites. The three remaining cell lines Co115, HT29, and SW48 displayed full methylation of all sites.

**Quantitative Gene Expression of \textit{ADAMTS1}, \textit{CRABP1}, and \textit{NR3C1} in Colon Cancer Cell Lines**

Figure 4 shows the relative expression levels of \textit{ADAMTS1}, \textit{CRABP1}, and \textit{NR3C1} in the colon cancer cell lines (n = 20). The expression levels are displayed as ratios between the individual genes and the endogenous control \textit{PGK1} and multiplied by a factor of 1000. The housekeeping genes \textit{PGK1} and \textit{ACTB} had overall comparable expression levels, but since the median standard deviation in \textit{PGK1} was less than in \textit{ACTB} (0.12 and 1.69, respectively), the gene expression of \textit{PGK1} was chosen for normalization of the
Figure 3. Methylation status of individual CpG sites by bisulphite sequencing in ADAMTS1 (A), CRABP1 (B), and NR3C1 (C). The upper part of each panel is a schematic presentation of the CpG island in the area of transcription start amplified by the bisulphite sequencing primers. The transcription start site is represented by +1 and the vertical bars indicate the location of individual CpG sites. The two arrows indicate the location of the methylation specific-PCR primers. For the middle part of each panel, filled circles represent methylated CpGs; open circles represent unmethylated CpGs; and open circles with a slash represent partially methylated sites (the presence of approximately 20-80% cytosine, in addition to thymine). The column of U, M and U/M at the right side of this middle part lists the methylation status of the cell line from MSP analyses of the respective genes. The lower part of each panel (highlighted in grey) is a section of the bisulphite sequence electropherogram. Abbreviations: MSP, methylation specific-PCR; s, sense; as, antisense; U, unmethylated; M, methylated; U/M, presence of both unmethylated and methylated band.
Figure 4. ADAMTS1 (A), CRABP1 (B), and NR3C1 (C) expression in colon cancer cell lines. The quantitative expression levels are displayed as ratios between the specified gene and the endogenous control PGK1 multiplied by a factor of 1000. The CpG island hypermethylation status assessed by methylation specific-PCR (MSP) is shown below each chart. Filled circles represent methylated CpGs; open circles with a slash represent presence of both methylated and unmethylated alleles. The cell lines are presented in the same order as in figure 2.
additionally close to zero. The two cell lines harboring both methylated and unmethylated DNA molecules from NR3C1 promoter methylation analyses revealed gene expression, although at various levels, as did the fully unmethylated cell lines.

**Discussion**

We have used microarray gene expression in combination with 5-aza-2’-deoxycytidine treatment of colon cancer cell lines to identify novel gene targets epigenetically inactivated in early stages of colorectal tumorigenesis. When cell lines are cultured with the DNA methyltransferase inhibitor 5-aza-2’-deoxycytidine, epigenetic gene silencing by aberrant DNA methylation can be reversed. This could be seen in the present study by the reappearance of unmethylated MSP bands of originally methylated control genes (data not shown). Genes up-regulated in cell lines after 5-aza-2’-deoxycytidine treatment might be potential candidates for inactivation by hypermethylation in the respective cancer type.

However, to increase the probability of selecting true epigenetic targets, we analyzed four individual colon cancer cell lines, including both MSI and MSS cell lines. Only array elements up-regulated two times or more in at least three of the four cell lines analyzed were investigated further.

Genome wide DNA microarray approaches has also previously been used to identify new target genes for epigenetic inactivation in colorectal cancer (22; 23) as well as in other cancer types (24; 25). Suzuki and co-workers elegantly used this approach culturing the colon cancer cell line RKO with trichostatin A and low-dose-5-aza-2´deoxycytidine. They identified a group of genes whose expression was minimally increased after 5-aza-2´deoxycytidine treatment alone, but was significantly induced by a combined treatment of both drugs. Twelve genes belonging to this group (including three known control genes) were shown to be fully methylated in the RKO cell line (22). Our gene expression data from the four drug-treated colon cancer cell lines includes information about eight of the twelve fully methylated genes identified by Suzuki *et al.* Five of these genes responded to our 5-aza-2´deoxycytidine treatment in one or more of the cell lines, whereas *FOLH1, PCDH8,* and *TIMP2* did not. As *FOLH1* and *PCDH8* were shown to be equally methylated in primary colorectal carcinomas and their normal counterparts (22), their relevance to tumorigenesis remains to be determined. *TIMP2* turned out to be methylated only in the RKO cell line and not in a panel of eight additional
colon cancer cell lines or in primary colorectal cancers (22). By only selecting genes that were induced in minimum three out of four cell lines upon drug treatment we have shown that the probability of identifying true epigenetic target genes with importance for tumorigenesis is increased. By subsequent investigation of the expression of the same genes in primary colorectal carcinomas (data not shown), we pinpointed the genes also down-regulated in the primaries, further improving the approach. After searching for CpG islands we finally selected ADAMTS1, CRABP1, NDRG1, and NR3C1 for methylation analyses as their encoded proteins have functions that could play a potential role in tumorigenesis.

Three of our four candidates were hypermethylated in several colon cancer cell lines by MSP analyses. As this method tends to overestimate DNA methylation in certain cases (26), we additionally bisulphite sequenced approximately half of the cell lines. For all genes, there seemed to be a good association between the results obtained with the two methods, which diminishes the rate of potential false positives from the MSP analyses. Rand et al have pointed out that incompletely converted sequences during bisulphite treatment can be coamplified with the methylated alleles during MSP (26). In the present study all cytosines at non-CpG sites in the bisulphite sequences were converted to thymine, and we could not detect any samples with incomplete conversion. Additionally, none of our MSP primer sets amplified untreated DNA (data not shown), which means that if unconverted sequences had been present, it would not have influenced the results. To establish whether the promoter hypermethylation of the present target genes led to transcriptional inactivation, we analyzed the quantitative gene expression in the cell lines. Seemingly, there was a good association between hypermethylated status and reduced gene expression, which was significant for two of the genes analyzed. The relatively small sample set might explain the lack of significance for ADAMTS1.

DNA from cell lines is often more frequently hypermethylated than is DNA from primary tumors (27). To investigate the potential role of ADAMTS1, CRABP1, and NR3C1 in colorectal tumorigenesis in vivo we analyzed a series of primary colorectal carcinomas and adenomas. As all three candidates displayed methylation in adenomas as well as primary tumors (although to various extents), the epigenetic inactivation of these novel target genes seems to be an early event in colon cancer development. CpG island hypermethylation has also previously been demonstrated to be an early and frequent
Epigenetic target genes in colorectal cancer

Event during colorectal carcinogenesis, but only a limited number of epigenetic changes have been identified (9; 28). Here, we add three novel genes to the list of epigenetically inactivated targets in the colorectal tumorigenesis.

ADAMTS1 is a metalloproteinase of the ADAM (29) family containing a thrombospondin type 1 motif and was first described in 1997 as an inflammation and cancer cachexia-related gene (30). Today, a total of 19 similar genes have been characterized in the human genome (29) and several of them have been implicated in various diseases, including cancer (31). The carboxyl-terminal half region of ADAMTS1 has been shown to suppress both tumorigenicity and experimental tumor metastatic potential (32) and by binding and sequestrating VEGF165, ADAMTS1 inhibits angiogenesis and endothelial proliferation (33). In the present study the CpG island of ADAMTS1 is frequently hypermethylated in colorectal adenomas, carcinomas, and cell lines, with subsequent decrease or loss of gene expression in the majority of methylated samples analyzed. Even though two of the cell lines methylated for ADAMTS1 had slightly higher gene expression levels than expected (TC71 and LS1034, see Figure 4), the expression was still three to four fold less than in the unmethylated cell line Colo320. The bisulphite sequences further revealed that both cell lines contained some partially methylated CpG sites. These sites were not overlapping, but located in close proximity to the transcription start site, which indicates that basically all CpG sites around the transcription start site of ADAMTS1 need to be fully methylated in order to completely silence the transcription. As the majority of the methylated cell lines fulfilled this criterion, we suggest that ADAMTS1 has an important role in colorectal tumorigenesis. Moreover, ADAMTS1 might be an epigenetic target of general importance in cancer as down-regulated gene expression of ADAMTS1 is found in invasive breast carcinoma (34), pancreatic cancer, and hepatocellular carcinoma (35).

The cellular retinoic acid binding proteins (CRABP) have a high affinity to retinoic acid and belong to a family of small cytosolic lipid binding proteins. Two highly homologous forms of CRABP have been characterized, CRABP1 and CRABP2, with about 75% amino acid identity (36). Although their exact function is not completely understood, regulation of the availability of retinoic acid to its nuclear receptors and subsequent enhancement of the retinoic acid effect on regulating target gene expression is among the plausible alternatives (37). The retinoic acid, a metabolite of vitamin A, plays an
important role in cell cycle arrest, cell differentiation, and to a certain extent apoptosis by altering gene transcription, and there is increasing evidence that retinoic acid metabolism may be altered during carcinogenesis (37). Whether or not these metabolic alterations are due to changes in retinoic binding proteins remains to be established, but several reports have documented alterations in the expression of retinol and retinoic acid binding proteins in various tumors types (37). In the present study we demonstrate that the inactivation of CRABP1 is present in subgroups of benign and malignant tumors of the large bowel. All cell lines with CRABP1 hypermethylation have reduced or no gene expression, which indicate that the promoter hypermethylation leads to transcriptional inactivation. This is also the case for the methylated cell lines displaying local partial methylation around CpG sites 6 to 11. Furthermore, CRABP1 is also known to be inactivated by promoter hypermethylation in papillary thyroid carcinomas and in the MSI colon cancer cell lines SW48 and HCT116 (38). Finally, the mouse Crabp1 gene, which reveal a 99.3% amino acid identity to human CRABP1 (36), demonstrates developmentally regulated gene expression through methylation changes in its 5′-flanking region (39). Taken together, these lines of evidence support that CRABP1 expression is epigenetically deregulated early in the establishment of a visible colorectal tumor.

The nuclear receptor subfamily 3, group C, member 1 (NR3C1) gene encodes the glucocorticoid receptor, which resides in the cytoplasm in a multiprotein complex. Upon binding to glucocorticoid the protein translocates into the nucleus where it functions as a transcription factor, and participates in the regulation of several molecular processes such as inflammation, cell growth, differentiation, and glucocorticoid-induced apoptosis (40). The ligand-activated glucocorticoid receptor is especially efficient in killing certain cells of the lymphoid lineage, and glucocorticoids are therefore included in essentially all chemotherapy protocols for lymphoid malignancies (41). Glucocorticoids can also repress cell cycle progression in a number of other cell types, including neoplastic thymic epithelial cells (40) and non-small cell lung cancer cells (42). In primary colon cancers, the immunohistochemical expression of NR3C1 has been correlated with the cell cycle-related molecules pRb, and p16 (43). Crosstalk between NR3C1 and the central tumor suppressor TP53 has also been indicated (44). The normal population has a variable sensitivity to glucocorticoids, which in part can be explained by genetic
changes reported in the \( NR3C1 \) gene, including several polymorphisms and some mutations (45). As the level of \( NR3C1 \) expression is a critical determinant for glucocorticoid sensitivity (46), the transcriptional inactivating promoter hypermethylation found in the present study in colorectal carcinomas and to a less degree in adenomas will render these neoplasms glucocorticoid resistant. To our knowledge, this is the first study reporting promoter hypermethylation of \( NR3C1 \) and subsequent loss of expression. If the inactivation of \( NR3C1 \) is confirmed also in other cancers, it might be of clinical importance in treatment of cancer patients receiving glucocorticoids, like lymphoma and acute lymphatic leukemia patients. Interestingly, microarray expression data of such samples reveal frequent down-regulation of \( NR3C1 \) relative to a common reference (47) (see SOURCE for the expression of \( NR3C1 \) in lymphomas: http://source.stanford.edu/cgi-bin/source/sourceSearch). Indeed, \( NR3C1 \) was down-regulated in blast cells from five of eight acute lymphoblastic leukemia patients with relapse compared with blast cells at diagnosis (see GDS363 report deposited in NCBI's Gene Expression Omnibus; GSE, http://www.ncbi.nlm.nih.gov/geo/).

In conclusion, by using strict selection criteria for identifying genes silenced by promoter hypermethylation we present three novel epigenetically inactivated genes of importance in colorectal cancer: \( ADAMTS1 \), \( CRABP1 \), and \( NR3C1 \).

References


38. Huang Y, de la CA, Pellegata NS. Hypermethylation, but not LOH, is associated with the low expression of MT1G and CRABP1 in papillary thyroid carcinoma. Int J Cancer 2003;104:735-744.


Correspondence: Ragnhild A. Lothe, Department of Genetics, Institute for Cancer Research, The Norwegian Radium Hospital, 0310 Oslo, Norway.
E-mail: rlothe@radium.uio.no Telephone number: +47 22934415, Fax number: +47 22934440.

The study was funded by a grant (A95068, RAL) from the Norwegian Cancer Society supporting GEL and KK as a Research Fellows. The authors are grateful for Dr. Mario Fraga’s expertise help with the primer design of the NR3C1 bisulphite sequencing primers.
DNA methylation in benign and malignant large bowel tumors and their in vitro models

Manuscript
In order to gain insights in the molecular evolution of epigenetics that parallels the development of colorectal cancer we examined the promoter methylation of eleven selected genes, including \textit{ADAMTS1}, \textit{APC}, \textit{CRABP1}, \textit{HOXA9}, \textit{MGMT}, \textit{MLH1}, \textit{NR3C1}, \textit{p16^{INK4A}}, \textit{PTEN}, \textit{RUNX3}, and \textit{SCGB3A1}, in hyperplastic polyps (n = 12), adenomas (n = 63), carcinomas (n = 53), and in colon cancer cell lines (n = 20).

In total, 83\% hyperplastic polyps, 73\% adenomas, 89\% carcinomas, and 100\% cell lines were methylated in one or more of the 11 genes analyzed with an average of 4.1, 1.7, 3.4, and 5.3 methylated genes in the four respective tumor groups. The most frequently methylated genes in the primary carcinomas were \textit{ADAMTS1}, \textit{CRABP1}, and \textit{MGMT}. Three genes, \textit{HOXA9}, \textit{RUNX3}, and \textit{SCGB3A1} are for the first time reported to be methylated in benign lesions of the large bowel, as well as \textit{ADAMTS1}, \textit{CRABP1}, and \textit{NR3C1} in hyperplastic polyps.

Methylation frequencies among the individual genes typically increased with malignancy, with the apparent exception of hyperplastic polyps. These lesions displayed methylation frequencies comparable to those of carcinomas and may belong to the subgroup of serrated adenomas. In addition, three genes, \textit{HOXA9}, \textit{MGMT}, and \textit{APC} showed similar methylation frequencies in adenomas and carcinomas, suggesting that the inactivation of these occur early in colorectal tumorigenesis. We confirmed that methylation is most common in carcinomas with MSI and proximal location. Additional associations to gender, age and polyp size were also identified for some of the methylated genes.

In general, methylation frequencies were higher in cell lines than in primary tumors and statistically significant for \textit{CRABP1}, \textit{p16^{INK4A}}, and \textit{SCGB3A1}. However, as the overall methylation profiles of the two groups were comparable, colon cancer cell lines can be considered representative epigenetic models for large bowel carcinomas.

The last decade, numerous reports have demonstrated the importance of epigenetic changes in human tumors, and cancer is now recognized to be an epigenetic – as well as a genetic - disease [1,2]. During the development of colorectal cancer, both genetic and epigenetic changes accumulate [3]. The majority (~85\%) of the sporadic carcinomas are characterized by several chromosomal aberrations, often
referred to as a chromosomal instability (CIN) phenotype [4]. A second route to colorectal cancer is the microsatellite instability (MSI) pathway. The MSI phenotype is characterized by multiple insertions and deletions of short nucleotide repeats in coding as well as in non-coding sequences throughout the genome [5]. This instability is caused by a defect DNA mismatch repair system and is found in a subgroup of colorectal carcinomas (~15%) [6-8]. Finally, a third pathway for colorectal cancer development has been suggested, the so-called CpG island methylator phenotype (CIMP) [9]. CIMP positive tumors display simultaneous methylation of several loci, are associated with proximal location in the colon, and overlap considerably with the MSI phenotype. However, the CIMP concept remains controversial [10,11].

Adenomas are known to be common precursors of colorectal carcinomas, and increasing size and amount of villous component parallels the malignant potential of these lesions [12]. Some genetic aberrations, such as APC and KRAS2 gene mutations, are typically found in adenomas [13-16], but the influence of epigenetic alterations is not well elucidated. Another type of colorectal polyps, the hyperplastic polyps have long been considered as benign lesions with no malignant potential, however, recent findings suggest that a subgroup of these lesions also can progress into malignancy [18-20]. These sessile serrated adenomas may give rise to both MSS and MSI carcinomas, depending on the nature of the accumulated molecular changes [17,21].

Cell lines are in general easy to culture and the unlimited amounts of DNA, RNA, and proteins make them well suited and commonly used as experimental tools. However, the in vivo relevance of tumor cell lines is debated, and methylation profiles of colorectal cancer cell lines remain mostly unknown.

In the present study the DNA promoter methylation have been examined in 11 selected genes, including known and novel methylation targets in cancer. In order to further address the influence of aberrant methylation during the development of colorectal tumors, we have compared DNA promoter methylation profiles among hyperplastic polyps, adenomas, carcinomas, and colon cancer cell lines. Finally, the promoter methylation status of individual genes was also examined in relation to known genetic and clinicopathological variables.

**Materials and methods**

**Tumors and cell lines**

Twelve colorectal hyperplastic polyps from 11 individuals, 63 adenomas from 52 individuals, 53 carcinomas from 52 patients, and 20 colon cancer cell lines derived from 18 individuals were included in the present study, along with three normal colon mucosa samples from individual colorectal cancer patients (derived in distance from the tumors). The hyperplastic polyps and adenomas were obtained from a Norwegian Polyp Study; samples were taken from volunteers participating in a polyp screening initiated at one hospital in south-east of Norway 1983 [22]. The carcinomas are from an unselected prospective series collected from 7 hospitals in the south-east region of Norway during 1987-1989 [23]. Genetic and clinicopathological variables, including MSI
status, sex, age at cancer diagnosis/age at polyp removal, tumor site, polyp size, and multiplicity, are evaluated in relation to the present data.

**Bisulphite modification, candidate genes, and methylation specific-PCR (MSP)**

DNA from all samples was bisulphite modified, a chemical process in which unmethylated but not methylated cytosines are converted to uracil [24-26]. The benign lesions were bisulphite modified using the CpGenome™ DNA modification kit (Serological, Norcross, GA, USA) following the manufacturers instructions. The colorectal carcinomas and colon cancer cell lines were treated as described previously [24,27].

Promoter methylation of **ADAMTS1, APC, CRABP1, HOXA9, MGMT, MLH1, NR3C1, p16<sup>INK4a</sup>, PTEN, RUNX3, and SCGB3A1** was analyzed by methylation specific-PCR (MSP) [28] in whole or parts of the various tumor series. The methylation status of **APC, MGMT, MLH1, and p16<sup>INK4a</sup>** in the carcinomas and cell lines [27], as well as the methylation status for **ADAMTS1, CRABP1, and NR3C1** in the adenomas, carcinomas and cell lines [29] have previously been published by us.

### Table 1. PCR primers used for methylation specific-PCR and microsatellite instability analyses.

<table>
<thead>
<tr>
<th>Primer set</th>
<th>Sense primer</th>
<th>Antisense primer</th>
<th>Frg-size bp.</th>
<th>An. Temp</th>
<th>MgCl₂ mM</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADAMTS1-M</strong></td>
<td>GTG AGT AAT ATC GTA GTT AAG GGC G</td>
<td>CTA AAA CAA AAA ACG TCT AAG ACG C</td>
<td>103</td>
<td>59</td>
<td>1.5</td>
<td>[29]</td>
</tr>
<tr>
<td><strong>ADAMTS1-U</strong></td>
<td>GTG AGT AAT ATG GTA GTT AAG G</td>
<td>AAA ACA AAA AACT CTA AAA AAC C</td>
<td>101</td>
<td>59</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>APC-M</strong></td>
<td>TAT TGC GTA GTG GGC GTC</td>
<td>TGG ACG AAc TCC GGA GCA</td>
<td>98</td>
<td>63</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td><strong>APC-U</strong></td>
<td>GTG TTT TAT TGT GGA GTG TGG GTT</td>
<td>CCA ATC AAC AAA CTC CCA ACA A</td>
<td>108</td>
<td>57</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td><strong>BAT-25</strong></td>
<td>TCG CCT CCA AGA ATG TAA GT</td>
<td>TCT GCA TTA TAA CT TAA TGG CTC</td>
<td>124</td>
<td>55</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>BAT-26</strong></td>
<td>TGA CTA CTT TGT AGT TCA GCC</td>
<td>AAC CAT TCA ACA TTT TTA ACC C</td>
<td>122</td>
<td>55</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>CRABP1-M</strong></td>
<td>GGA GGT TTT TTA GTT GGA GAG C C</td>
<td>CTC GCA AAA CGA AAA ACA AGC</td>
<td>211</td>
<td>57</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>CRABP1-U</strong></td>
<td>GAG GTT TTT TAG TTAG GAG AGT G</td>
<td>AAC TCA CAA AAC AAA AAC TAA CAC T</td>
<td>212</td>
<td>57</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>HOXA9-M</strong></td>
<td>GGT TAA TGG GGG GCC G GG GTT C</td>
<td>AAG GCC TAA CCC GCC CCA CCG G</td>
<td>127</td>
<td>68</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>HOXA9-U</strong></td>
<td>GTA TGG TTA ATG GGG GTT TGG GTT TT</td>
<td>CCA TAC CCA AAT TCC CCT ACC CCC</td>
<td>139</td>
<td>66</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>MGMT-M</strong></td>
<td>TTT CGA GTG TAC TAG GTT GCC</td>
<td>GCA ATC TCA CGA AAA ACA AAG C</td>
<td>81</td>
<td>58</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td><strong>MGMT-U</strong></td>
<td>TTT GTG TTT TGT TGT TGT TGT TT</td>
<td>AAT TCC AAC ATT TCC CCA AAA AAA CAA A</td>
<td>93</td>
<td>58</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td><strong>MLH1-M</strong></td>
<td>CCG ATAG CCG ATT TTA ACG TGA TAC</td>
<td>CCG TCC TCC CTA AAA CG</td>
<td>72</td>
<td>53</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>MLH1-U</strong></td>
<td>TGG ATA GTG ATT TTA ATT ATG TTA AAG C</td>
<td>ATC CCT CCA TAA AAC AAT TAC TAC CCA</td>
<td>71</td>
<td>60</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td><strong>NR3C1-M</strong></td>
<td>TGT TGG TTT TGT TTT TGT TTT GT</td>
<td>CCG CCC GAT CCC AAC TAC TTC GAC</td>
<td>196</td>
<td>69</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>NR3C1-U</strong></td>
<td>TTA GTG TGG TTT TGT TGT TTA TAC</td>
<td>CCA TCC CAA CAA CTA CTA CTT CAA C</td>
<td>197</td>
<td>61</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>p16&lt;sup&gt;INK4a&lt;/sup&gt;-M</strong></td>
<td>TTA TTA GAG GGG GGG CAT GGC</td>
<td>GAC CCC CAA CAG CCA CCG TAA</td>
<td>150</td>
<td>63</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>p16&lt;sup&gt;INK4a&lt;/sup&gt;-U</strong></td>
<td>TTA TTA GAG GGT GGG GTG CAT T</td>
<td>CAA CCC CAA CAA TAA</td>
<td>151</td>
<td>64</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>PTEN-M</strong></td>
<td>GCC GCC GTG GCC GGT T</td>
<td>GAC TCC CGG AAA ACG CTA C</td>
<td>71</td>
<td>65</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>PTEN-U</strong></td>
<td>GAG AGA TGG TGG TGG TTG T</td>
<td>AAC TCC CAA AAA ACA CTA CC</td>
<td>78</td>
<td>61</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>RUNX3-M</strong></td>
<td>TTA CGA GGG GCC GTG GTC CAA GGG</td>
<td>AAA ACC ACC GAC CCG AGC TCC TCC</td>
<td>250</td>
<td>67</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>RUNX3-U</strong></td>
<td>TTA TTA GAG TGG GTG TTA TGT GGG</td>
<td>AAA AAC ACC AAC ACAA AAC ACC TCC</td>
<td>240</td>
<td>65</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>SCGB3A1-M</strong></td>
<td>GGT AGG GTG TTT TTA CCG TTC GTC</td>
<td>AAC TTA TAA TCC CCG ATC CTC G</td>
<td>135</td>
<td>61</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td><strong>SCGB3A1-U</strong></td>
<td>GGT AGG GTG TTT TTA CCG TTC G</td>
<td>CAA AAC TCC TAA TCC CCA ATC CTC C</td>
<td>135</td>
<td>61</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Frg-size bp.; Fragment size in base pairs, An. Temp; annealing temperature in degrees Celcius.

---

[a) http://www.ensemble.org/
Two primer sets are used for MSP, the first recognizes and anneals to methylated sequences only, whereas the second set amplifies unmethylated alleles. The unmethylated and methylated reactions of MSP were carried out in a total volume of 25 µl containing 1 x PCR Buffer (QIAGEN Inc., Valencia, CA), 200 µM dNTP (Amersham Pharmacia Biotech Products Inc., Piscataway, NJ), 20 pmol of each primer, and 0.625 U HotStarTaq DNA Polymerase (QIAGEN). All primers were purchased from Medprobe AS (Oslo, Norway) and the sequences are listed in Table 1 along with the magnesium concentration in each PCR reaction, PCR program annealing temperature, and fragment size.

Human placental DNA (Sigma Chemical Co., St. Louis, MO) treated in vitro with SssI methyltransferase (New England Biolabs Inc., Beverly, MA) was used as a positive control for MSP of methylated alleles, whereas DNA from normal lymphocytes was used as a control for unmethylated alleles. Water was used as a negative PCR control in both reactions.

Threshold for scoring methylated samples from methylation specific-PCR

The amount of PCR product generated by the methylation-specific primer set varied among the different samples. The band intensity of each sample was compared with that of the positive control and visually scored as previously described [30]. In brief, samples scored as + have bands that are less intense than the positive control in the methylation-specific reaction, whereas samples denoted ++ have equal or stronger band intensity than the positive control. For primary carcinomas and cell lines, only samples scored as ++ are recorded as methylated. For the benign lesions, both + and ++ samples are recorded as methylated. Two of the authors independently scored all samples and the methylation status of all positive samples was confirmed by a second independent round of MSP as well as a second scoring of both researchers.

Assessment of MSI status

To determine the MSI status of the benign tumors, two mononucleotide repeats, BAT-25 and BAT-26, were analyzed in tumor samples from each individual. The combined analyses of these two markers can detect more than 99% of the tumors with MSI [31]. Fifty ng DNA template was amplified in a 10 µl reaction volume consisting of 1xQIAGEN® PCR buffer containing 1.5mM MgCl₂ (QIAGEN GmbH, Hilden, Germany), 1.5-4 pmol BAT-25 primers (sense primer labeled with HEX in the 5’end), 1.5-4 pmol BAT-26 primers (sense primer labeled with 6-FAM in the 5’ end; DNA Technology AS, Aarhus, Denmark), 2 pmol of each of the four dNTPs, and 0.4 units of HotStarTaq™ Polymerase (QIAGEN®). The PCR reaction annealing temperature was 55°C and the program included 27 cycles. Primer sequences are listed in Table 1.

One µl PCR product was mixed with 0.5µl GeneScan®-500 [TAMRA] Size Standard (PE Biosystems) and 12µl deionized formamide (Kodak Eastman Chemical Company, New Haven, CT, USA), denatured and separated by 16 minutes capillary electrophoresis at 60°C (ABI PRISM™310 Genetic Analyzer; PE Biosystems). The results were independently scored by two authors. A second round of
analyses confirmed the MSI status. The MSI status was known for all carcinomas determined by use of the two BAT markers as well as by several dinucleotides [32].

Statistics
Two x 2 contingency tables were analyzed using the Fisher’s exact test. Three x 2 tables were analyzed by the Pearson $\chi^2$ test. Non-parametric analyses were performed using the Kruskal Wallis – and Mann-Whitney - tests. All $P$ values are derived from two tailed statistical tests using the SPSS 12.0 software (SPSS, Chicago, IL, USA) and $P < 0.05$ were considered statistically significant.

In the present study, multiple statistical tests have been performed to investigate the relationship between promoter DNA methylation and several clinicopathological variables (Table 2). When performing many tests, the likelihood of false positives increases and caution should therefore be taken when interpreting the results. [33].

Results
DNA hypermethylation in colorectal polyps, carcinomas, and colon cancer cell lines
The results of the MSP analyses of benign- and malignant tumors and of colon cancer cell lines are summarized in Figure 1 and Table 2. Detailed information about the methylation status of all 11 genes in each colon cancer cell lines is presented in Table 3, and representative raw data can be seen in Figure 2.

In total, 10/12 (83%) hyperplastic polyps, 46/63 (73%) adenomas, 47/53 (89%) carcinomas, and 20/20 (100%) cell lines were methylated in one or more of the 11 genes analyzed. The average number of genes methylated per sample was 1.7 for adenomas, 4.1 for hyperplastic polyps, 3.4 for carcinomas, and 5.3 for cell lines, and was significantly different among the groups ($P = 0.006$; median 1.0, 5.5, 3.0, and 5.5, respectively). Overall, hyperplastic polyps and carcinomas displayed more methylation than did adenomas ($P = 0.006$ and $P = 0.001$, respectively), whereas cell lines displayed more methylation than did
adenomas, whereas the hyperplastic polyp mean was 7mm. The mean age of individuals at polyp removal was 67 years. Eight mm is the mean size of cancer diagnosis was 68 years. The mean age of patients at time of listing for individual polyps. Below the line the patients’ sex, age, and number of polyps (multiplicity) are listed. Regarding multiplicity, the colon is defined to be methylated if one or more of the polyps are methylated. The mean age of patients at time of polyp localization, size, and MSI status data are entered and subdivided into benign polyps (adenomas) and adenocarcinomas (P = 0.007). Statistically significant differences in methylation frequencies among tumor groups were also evident at the single gene level. These differences are summarized in Figure 3, and the genes are classified into two categories according to their methylation frequencies among the various tumor subgroups. Genes (ADAMTS1, CRABP1, MLH1, NR3C1, p16INK4a, RUNX3, and SCGB3A1) showing increasing methylation frequencies from adenomas, to carcinomas (Fig.3a); and genes (APC, HOXA9, MGMT) displaying overall equal methylation frequencies in all tumor subgroups (Fig.3b). PTEN was unmethylated in all colon cancer cell lines as well as in carcinomas, and thus was not investigated in the polyps. The mean age of patients at time of cancer diagnosis was 68 years. The mean age at polyp removal was 67 years. Eight mm is the mean size of adenomas, whereas the hyperplastic polyp mean was 7mm. Table 2. CpG island methylation of selected genes compared with the patients clinicopathological features and tumor genotypes. Abbreviations: NS, not significant; Gen. Clin. and Path. Features, Genetic, Clinical and Pathological Features; HPs, hyperplastic polyps; MSI, microsatellite instable; MSS, microsatellite stable. Polyp localization, size, and MSI status data are entered and subdivided into benign polyps (adenomas) and adenocarcinomas (P = 0.007). Statistically significant differences in methylation frequencies among tumor groups were also evident at the single gene level. These differences are summarized in Figure 3, and the genes are classified into two categories according to their methylation frequencies among the various tumor subgroups. Genes (ADAMTS1, CRABP1, MLH1, NR3C1, p16INK4a, RUNX3, and SCGB3A1) showing increasing methylation frequencies from adenomas, to carcinomas (Fig.3a); and genes (APC, HOXA9, MGMT) displaying overall equal methylation frequencies in all tumor subgroups (Fig.3b). PTEN was unmethylated in all colon cancer cell lines as well as in carcinomas, and thus was not investigated in the polyps. PTEN is therefore not included in the figures, tables or statistics.
All 11 genes submitted to methylation analyses were unmethylated in the three normal colon tissue samples included in the present study.

**MSI status of benign tumors**

Three of seventy-five (4%) polyps displayed a microsatellite unstable phenotype, including one small (4mm) hyperplastic polyp and two large (>10mm) adenomas. Two of the MSI positive samples were located in the right side of the colon, whereas the last was located in the rectum.

**Associations among DNA methylation and genetic – and - clinicopathological characteristics**

The DNA methylation status compared with the genetic and clinicopathological features of the benign and malignant tumors are summarized in Table 2. Overall, gene methylation frequencies were higher in MSI than in MSS carcinomas, and statistically significant for **CRABP1, MLH1, NR3C1, RUNX3, and SCGB3A1** ($P < 0.0001$, $P < 0.0001$, $P = 0.002$, $P < 0.0001$, and $P = 0.006$, respectively). The same trend was seen for the MSI versus MSS cell lines, although not statistically significant.

There was a strong association between DNA methylation and proximal tumor location. Statistically significant for the following genes in carcinomas (**CRABP1, MLH1, NR3C1, RUNX3, and SCGB3A1**; $P < 0.0001$, $P < 0.0001$, $P = 0.001$, $P = 0.001$, and $P = 0.037$, respectively), and to a certain extent in hyperplastic polyps (**CRABP1**; $P = 0.055$). It should be noted that only 2/11 hyperplastic polyps were removed from the left side of the colon, the remaining were right-sided. Adenomas represented the only group in which a tendency towards more methylation in samples from the distal colon (rectum) than in samples from the proximal colon could be seen and then only for one gene (**HOXA9**, $P = 0.078$).

---

**Table 3. Promoter methylation status of colon cancer cell lines.** The 20 cell lines were stratified according to their MSI status, and analyzed by methylation specific-PCR (MSP).

Abbreviations: MSI, microsatellite unstable; MSS, microsatellite stable; U, unmethylated; M, methylated; U/M, partially methylated (the presence of both methylated and unmethylated alleles). The methylation status for 7 out of 11 genes has previously been published in the same cell lines [21] [24].
Adenomas of equal or larger size than the mean value (8 mm in the present study) seem to have higher methylation frequencies than do smaller adenomas. Although this was only statistically significant for \textit{ADAMTS1} ($P = 0.020$), the same trend is seen for \textit{HOXA9} and \textit{p16}^{\text{INK4a}} ($P = 0.061$ and $P = 0.092$, respectively). For the hyperplastic polyps, no association could be found between methylation and size. The same test was performed using the more common 10mm as a cut-off, with comparable results.

Among the carcinomas, \textit{MLH1} and \textit{NR3C1} were more frequently methylated in women than in men ($P = 0.040$ and $P = 0.024$, respectively). \textit{RUNX3} display the same tendency among the malignant tumors ($P = 0.085$). However, all four methylated \textit{RUNX3} adenomas are derived from males ($P = 0.023$).

In the present study, we find no significant associations between the individuals’ age at polyp removal and DNA methylation. However,
both MLH1 and p16(INK4a) tend to show more methylation in carcinomas derived from patients of higher age ($P = 0.092$ and $P = 0.076$, respectively).

No statistically significant association between gene-specific methylation and polyp multiplicity were found

**Widespread methylation**

Several samples from each tumor group harbored simultaneous DNA methylation of two or more of the target genes analyzed (Figure 4). The distribution of methylated gene number per sample did not appear to be bimodal. Five of 11 (45%) hyperplastic polyps displayed methylation of five or more genes, whereas only six of 63 (10%) adenomas did the same ($P = 0.006$). In carcinomas, widespread methylation was seen more frequently in MSI- (14/28; 50%) than in MSS - (3/25; 12%) samples ($P = 0.046$). The same trend could be seen among the cell lines (Figure 4), although this was not statistically significant. The few adenomas displaying widespread methylation were by far larger in size (median = 21mm) than the rest (median = 8mm; $P = 0.038$).

**Discussion**

The present data demonstrate aberrant methylation of all analyzed genes, with the apparent exception of PTEN, in benign and malignant tumors as well as in their in vitro models. CpG island hypermethylation of gene promoters is associated with reduced or silenced expression of the gene in question, and can be partially relieved by demethylation of the promoter region [2,34,35]. This type of gene inactivation is acknowledged as a frequent mechanism in cancer development. The link between methylation and lack of expression has previously been confirmed, by us and others, for all the genes included in this study [28,29,36-40]. Therefore, the observed methylation is interpreted to have functional
consequence and to contribute to the tumorigenesis. Therefore, we suggest that the observed methylation has a functional consequence and contributes to the tumorigenesis.

In general, the methylation profiles of adenomas, carcinomas, and colon cancer cell lines demonstrate a stepwise increase in CpG island promoter methylation frequency with malignancy. This was evident both from a group-wise comparison using a Mann-Whitney test, and also at the single gene level. The apparent exception is the hyperplastic polyps, which displayed methylation frequencies comparable to those of the carcinomas. The hyperplastic polyps have originally been regarded as non-neoplastic lesions [21,41]. However, some years ago, a subgroup of hyperplastic polyps with malignant potential was identified [20]. These tumors, named sessile serrated adenomas [20], are typically larger than the true hyperplastic polyps, arise in the proximal colon, and show a high mutation rate in the \( \text{BRAF} \) gene [21]. They are suggested to be the precursor of MSI-cancers and have also been reported to harbor extensive DNA methylation [17]. The majority of hyperplastic polyps included in the present study were located in the right side of the colon and six of twelve also harbored a mutation in the \( \text{BRAF} \) gene (V599E; data not shown), indicating that some of the samples might belong to the sessile adenoma group.

For individual genes a higher methylation frequency is typically seen in carcinomas compared to adenomas, although this was not the case for \( \text{APC} \), \( \text{MGMT} \), and \( \text{HOXA9} \). These three genes display comparable methylation frequencies between the two tumor groups, indicating that they are early events in the tumorigenesis. Indeed, \( \text{APC} \) mutations have been reported to be present in aberrant crypt foci in the colon as well as in adenomas and carcinomas [42], and inactivating \( \text{APC} \) promoter hypermethylation has also been found in benign and malignant colorectal tumors [14,15,38,43-46]. The \( \text{APC} \) protein is part of a cytoplasmic protein complex regulating the balance of \( \beta \)-catenin (\( \text{CTNNB1} \)) degradation in the canonical WNT signaling pathway, and mutations and/or methylation of the \( \text{APC} \) gene will lead to accumulation of \( \beta \)-catenin and thereby increased transcription of downstream genes [47,48]. The \( \text{MGMT} \) gene is methylated in both MSS and MSI carcinomas, indicating that this event occurs prior to the separation of these two molecular pathways of carcinogenesis. This is supported by the fact that \( \text{MGMT} \) methylation has been identified even in the aberrant crypt foci [49]. The \( \text{MGMT} \) gene encodes a direct repair enzyme, which removes mutagenic and cytotoxic adducts from O\(^6\)-guanine in the DNA [50] and has previously been suggested – in concert with \( \text{BRAF} \) and \( \text{KRAS2} \) mutations - to be important in the transformation of benign cells into MSI-low and MSS colorectal tumors [17]. The third gene potentially inactivated early in the colorectal tumorigenesis, \( \text{HOXA9} \), is more often methylated in MSS than in MSI tumors and cell lines. This is in contrast with the majority of genes analyzed for promoter hypermethylation in colorectal carcinomas, and may point towards the importance of \( \text{HOXA9} \) inactivation in the initiation of the MSS pathway. To the best of our knowledge, this is the first report of \( \text{HOXA9} \) methylation in colorectal neoplasms. However, \( \text{HOXA9} \) methylation has previously
been associated with mortality in noninfant neuroblastoma patients [36] and has also been reported in testicular cancer [27,51]. \textit{HOXA9} belongs to the homeobox gene family, which are transcriptional master switches, regulating embryonic development [52]. Homeobox genes have also been shown to play an important role in various cancer types, such as leukemia and lung cancer, and can affect proliferation, differentiation, and apoptosis [52].

There are both technical and biological aspects that may influence the interpretation of gene methylation results. First of all, we have used two different thresholds to acknowledge a sample as methylated, assessed by the MSP assay, in benign compared to malignant tumors. Although the amount of PCR product generated by the methylation-specific primer set, and judged by gel band intensity, varies considerably between samples, it is highly reproducible for individual samples and may thus reflect the number of methylated alleles present. The primary carcinomas included in the present study are from a series evaluated to contain a mean number of 84% tumor cells [23]. Even though a minor fraction of these cells should display promoter methylation of the gene in question, it is highly unlikely that this will affect the carcinoma phenotype, since the majority of tumor cells will still produce the protein. Hence, only carcinomas and cell lines displaying strong gel bands from amplification with the methylation-specific primer set are acknowledged to be hypermethylated. Benign lesions on the other hand are expected to contain a mixture of cells. The admixture of unmethylated DNA from these cells will dilute the neoplastic epithelial DNA and thereby mask the true methylation status. We therefore acknowledge benign tumors with weak - as well as benign tumors with strong – methylation-specific gel bands to be methylated.

The less restrictive criterion for acknowledging methylated samples among the benign lesions might provide some false positive results. However, the methylation frequencies reported here for the six genes previously analyzed in colorectal lesions (\textit{APC}, \textit{MLH1}, \textit{MGMT}, \textit{p16\textsuperscript{INK4a}}, \textit{RUNX3}, and \textit{SCGB3A}) are overall comparable with other reports [11,27,38,39,43,44,53-59]. However, individual variations can be seen, and \textit{APC} methylation reported for adenomas is one example, in which the frequency ranges from 18% [38] to 60% [60]. A range of factors can cause this variation in methylation frequencies for a particular gene and cancer type (listed in [27]), including different inter-laboratory routines. In the present study, all genes have been analyzed in the same laboratory using large series of benign as well as malignant tumors and colon cancer cell lines. It should be noted that in general, most genes have weaker PCR products in the benign lesions compared with those of the carcinomas and cell lines.

Another important issue of DNA methylation analyses is the design of correct primers. Promoter hypermethylation of \textit{PTEN} has been frequently reported in various tumor types [63-66], including in colorectal cancer [67]. However, the majority of MSP primer sets used for these analyses have failed to discriminate between \textit{PTEN} and its frequently methylated pseudogene, leading to a high rate of false positives [68]. In the present study, we used MSP primers specifically designed to amplify
the protein encoding *PTEN* gene [69], and show that *PTEN* is not subjected to promoter hypermethylation in colorectal carcinomas nor in colon cancer cell lines. Previous reports have shown that colorectal polyps also are unmethylated for this gene [44,68,70]. The methylated products amplified by the remaining MSP primer sets used in the present study have all shown associations with lost or reduced protein expression (see references in table 1).

The present study confirms that methylation of *MLH1* is characteristic of right-sided sporadic colon tumors with microsatellite-instability [8,71-75]. Several of the other genes analyzed here (*CRABP1*, *NR3C1*, *RUNX3*, and *SCGB3A1*) were also more commonly methylated in MSI tumors than in MSS tumors, further supporting the hypothesis that DNA methylation plays a more important role in proximal than in distal carcinogenesis. Even though these associations were only statistically significant for the carcinomas, the same trend could be seen among the adenomas, as well as the hyperplastic polyps. Several (*n* = 9) of the individuals included in our polyp material had multiple polyps in the colon. To exclude potential bias in the material due to similarities between polyps from the same colon, one polyp from each individual was randomly selected for statistical analyses of the potential associations between methylation and tumor site. In large, the results were comparable.

Studies of *APC* and *KRAS2* have suggested that mutation frequencies and malignancy augment with increasing polyp size [76-78]. Here, we demonstrate that *ADAMTS1*, an inhibitor of both angiogenesis and endothelial proliferation [79], is more often methylated in large than in small adenomas. The same tendency could also be seen for *HOXA9* and *p16INK4A*, although this was not statistically significant. Further, the present results confirmed the gender differences previously reported for *MLH1* methylation [80,81]. Also *NR3C1*, encoding the glucocorticoid receptor, displayed more methylation among carcinomas derived from female than male patients. *RUNX3* on the other hand, seemed to be more frequently methylated in male than in female adenomas. However, among carcinomas, the methylation had a similar distribution as that of *MLH1* and *NR3C1*. Regarding age, *MLH1* and *p16INK4A* seemed to be more frequently methylated in lesions from patients of high age, than among younger patients. Other studies have also shown an increase in cancer specific methylation with greater age [82-85].

Previous studies have observed a concordant hypermethylation of multiple CpG islands in large bowel tumors, which has led to the proposition of a third pathway to colorectal tumorigenesis, the CpG island methylator phenotype (CIMP) [9,86]. In the original study, the CIMP negative samples only rarely harbored methylated gene promoters, resulting in a bimodal distribution [9]. This has been difficult to reproduce by others [11,43], initiating a debate regarding the existence of this third transforming pathway. The data presented here do not support CIMP, as the distribution of methylated promoters per sample seems to be continuous rather than bimodal in polyps and carcinomas as well as in the cell lines. However, we confirm that CpG island methylation is a frequent phenomenon...
in colorectal tumorigenesis and especially among MSI tumors.
Colon cancer cell lines can in many ways be regarded as in vitro models for the in vivo situation. Both gene changes [87] and genomic [88] aberrations have been extensively studied in these models and the results show that the colon cancer cell lines are representative of in vivo carcinomas. At the epigenetic level, cell lines are in general more methylated than are primary tumors [89]. However, among various cancer types examined by restriction landmark genome scanning (RLGS), cell lines originating from the large bowel have been shown to be the ones that most resemble their primary tumor counterpart [89]. The overall CpG island promoter methylation analyzed here increases significantly from carcinomas to colon cancer cell lines. However, at the single gene level, this is only evident for CRABP1 and SCGB3A1 in addition to the previously reported p16\(^{INK4a}\) [27]. Since the overall methylation profile is similar among the carcinomas and colon cancer cell lines (see figure 1), we conclude that the cell lines can be considered relevant epigenetic models.

Conclusions
Aberrant CpG island hypermethylation increases with malignancy, with the apparent exception of hyperplastic polyps. These lesions display methylation frequencies comparable with those of carcinomas and may belong to the subgroup of sessile serrated adenomas. Here, we show that colon cancer cell lines are typically more methylated than are primary tumors. However, the overall methylation profile of these two groups is comparable, indicating that the cell lines can be considered as representative epigenetic models for large bowel carcinomas. We also suggest that methylation of the homeobox gene HOXA9 is an early event in colorectal tumorigenesis, along with APC- and MGMT- methylation, previously reported by others.

Acknowledgements
The study was funded by grants from the Norwegian Research Council (161448/V40, RAL) and the Norwegian Cancer Society (A95068, RAL) supporting TA and GEL as Research Fellows, respectively.

Authors’ affiliations
T. Ahlquist, G.E.Lind, R.A.Lothe, Group of Molecular Cancer Genetics, Dept. of Genetics, Inst. for Cancer Research, The Norwegian Radium Hospital, Norway
E.Thiis-Evensen, Morten Vatn, Medical Department, National University Hospital, Oslo, Norway
G.I.Meling, Surgical Department, Faculty Division Akershus University Hospital, University of Oslo, Norway
T.O.Rognum, Institute of Forensic Medicine, The National Hospital, University of Oslo, Norway

Reference List


international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998, **58**: 5248-5257.


Methylation in Colorectal Neoplasia


81. **Malkhosyan SR, Yamamoto H, Piao Z, Peruchó M:** Late onset and high incidence of colon cancer of the mutator phenotype with hypermethylated hMLH1 gene in women. *Gastroenterology* 2000, 119: 598.


APPENDICES
Appendix I. General abbreviations

A      adenine
ACF    aberrant crypt foci
ATP    adenosine triphosphate
bp     base pair
C      cytosine
ChIP   chromatin immunoprecipitation
CIN    chromosome instability
CIMP   CpG island methylator phenotype
CpG    cytosine phosphate guanine
DNA    deoxyribonucleic acid
DNMT   DNA methyltransferase
FAP    familial adenomatous polyposis
G      guanine
HNPCC  hereditary nonpolyposis colorectal cancer
i(12p) isochromosome 12p
ICF    immunodeficiency, centromeric instability, facial anomalies
ITGCN  intratubular germ cell neoplasia
MSI    microsatellite instability
MSP    methylation specific-PCR
MSS    microsatellite stable
PGC    primordial germ cell
PCR    polymerase chain reaction
ref    reference
RLGS   restriction landmark genome scanning
RNA    ribonucleic acid
RT-PCR reverse transcription-PCR
SAM    S-adenosyl-L-methionine
T      thymine
TGCT   testicular germ cell tumor
WHO    world health organization
### Appendix II. Gene symbols

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Aliases</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ABCB1</em></td>
<td>ATP-binding cassette, sub-family B (MDR/TAP), member 1</td>
<td>MDR1, P-gp, CD243, GP170, ABC20</td>
</tr>
<tr>
<td><em>ADAMTS1</em></td>
<td>a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 1</td>
<td>C3-C5, METH1, KIAA1346</td>
</tr>
<tr>
<td><em>APAF1</em></td>
<td>apoptotic protease activating factor</td>
<td>CED4</td>
</tr>
<tr>
<td><em>APC</em></td>
<td>adenomatosis polyposis coli</td>
<td>DP2, DP3, DP2.5</td>
</tr>
<tr>
<td><em>AR</em></td>
<td>androgen receptor</td>
<td>DP2, DP3, DP2.5</td>
</tr>
<tr>
<td><em>BAX</em></td>
<td>BCL2-associated X protein</td>
<td></td>
</tr>
<tr>
<td><em>BCL2</em></td>
<td>B-cell CLL/lymphoma 2</td>
<td>Bcl-2</td>
</tr>
<tr>
<td><em>BNIP3</em></td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3</td>
<td>Nip3</td>
</tr>
<tr>
<td><em>BRCA1</em></td>
<td>breast cancer 1, early onset</td>
<td>RNF53</td>
</tr>
<tr>
<td><em>BTG1</em></td>
<td>B-cell translocation gene 1, anti-proliferative</td>
<td></td>
</tr>
<tr>
<td><em>CASP5</em></td>
<td>caspase 5, apoptosis-related cysteine protease</td>
<td>ICE(rel)III</td>
</tr>
<tr>
<td><em>CASP8</em></td>
<td>caspase 8, apoptosis-related cysteine protease</td>
<td>MCH5, MACH, FLICE</td>
</tr>
<tr>
<td><em>CCND2</em></td>
<td>cyclin D2</td>
<td></td>
</tr>
<tr>
<td><em>CD44</em></td>
<td>CD44 antigen</td>
<td>IN, MC56, Pgp1, MIC4, MDU2, MDU3, CD44R</td>
</tr>
<tr>
<td><em>CDH1</em></td>
<td>cadherin 1, type 1, E-cadherin (epithelial)</td>
<td>uvomorulin</td>
</tr>
<tr>
<td><em>CDH4</em></td>
<td>cadherin 4, type 1, R-cadherin (retinal)</td>
<td></td>
</tr>
<tr>
<td><em>CDH13</em></td>
<td>cadherin 13, H-cadherin (heart)</td>
<td>CDHH</td>
</tr>
<tr>
<td><em>CDKN1A</em></td>
<td>cyclin-dependent kinase inhibitor 1A (p21, Cip1)</td>
<td>P21, CIP1, WAF1, SDI1, CAP20, p21CIP1</td>
</tr>
<tr>
<td><em>CDKN2A</em></td>
<td>cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4), alpha transcript</td>
<td>CDK4l, p16, INK4a, MTS1, CMM2, ARF, p19, p14, INK4, p16INK4a</td>
</tr>
<tr>
<td><em>CDKN2A</em></td>
<td>cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4), beta transcript</td>
<td>CDK4l, p16, INK4a, MTS1, CMM2, ARF, p19, p14, INK4, p16INK4a</td>
</tr>
<tr>
<td><em>CDKN2B</em></td>
<td>cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)</td>
<td>P15, MTS2, INK4B</td>
</tr>
<tr>
<td><em>CDX1</em></td>
<td>caudal type homeo box transcription factor 1</td>
<td></td>
</tr>
<tr>
<td><em>CDX2</em></td>
<td>caudal type homeo box transcription factor 2</td>
<td></td>
</tr>
<tr>
<td><em>CGGBP1</em></td>
<td>CGG triplet repeat binding protein</td>
<td>p20-CGGBP</td>
</tr>
<tr>
<td><em>CGRRF1</em></td>
<td>cell growth regulator with ring finger domain</td>
<td>CGR19</td>
</tr>
<tr>
<td><em>CLDN4</em></td>
<td>claudin 4</td>
<td>CPETR, CPE-R, WBSCR8, hCPE-R</td>
</tr>
<tr>
<td><em>COL1A2</em></td>
<td>collagen, type I, alpha 2</td>
<td></td>
</tr>
<tr>
<td><em>CRABP1</em></td>
<td>cellular retinoic acid binding protein 1</td>
<td>CRABP</td>
</tr>
<tr>
<td><em>CTNNB1</em></td>
<td>catenin (cadherin-associated protein), beta 1, 88kDa</td>
<td></td>
</tr>
<tr>
<td><em>CXX1</em></td>
<td>CAAX box 1</td>
<td></td>
</tr>
<tr>
<td><em>DAB2IP</em></td>
<td>DAB2 interacting protein</td>
<td>AF9Q34, DIP1/2, KIAA1743</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>DAPK1</td>
<td>death-associated protein kinase 1</td>
<td>DAPK</td>
</tr>
<tr>
<td>DCC</td>
<td>deleted in colorectal carcinoma</td>
<td></td>
</tr>
<tr>
<td>DLX6</td>
<td>distal-less homeobox 6</td>
<td></td>
</tr>
<tr>
<td>DNAJC15</td>
<td>DnaJ (Hsp40) homolog, subfamily C, member 15</td>
<td>MCJ</td>
</tr>
<tr>
<td>DNMT1</td>
<td>DNA (cytosine-5-) methyltransferase 1</td>
<td>MCMT, CXXC9</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>DNA (cytosine-5-) methyltransferase 3 alpha</td>
<td></td>
</tr>
<tr>
<td>DNMT3B</td>
<td>DNA (cytosine-5-) methyltransferase 3 beta</td>
<td></td>
</tr>
<tr>
<td>ELK1</td>
<td>ELK1, member of ETS oncogene family</td>
<td></td>
</tr>
<tr>
<td>EMX2</td>
<td>empty spiracles homolog 2 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>EPHA7</td>
<td>EPH receptor A7</td>
<td>Hek11</td>
</tr>
<tr>
<td>ESR1</td>
<td>estrogen receptor 1</td>
<td>ER, NR3A1, Era</td>
</tr>
<tr>
<td>FANCF</td>
<td>Fanconi anemia, complementation group F</td>
<td>FAF</td>
</tr>
<tr>
<td>FAS</td>
<td>Fas (TNF receptor superfamily, member 6)</td>
<td>CD95, APO-1</td>
</tr>
<tr>
<td>FHIT</td>
<td>fragile histidine triad gene</td>
<td>FRA3B, AP3Aase</td>
</tr>
<tr>
<td>GAGE</td>
<td>G antigen</td>
<td></td>
</tr>
<tr>
<td>GATA4</td>
<td>GATA binding protein 4</td>
<td></td>
</tr>
<tr>
<td>GATA5</td>
<td>GATA binding protein 5</td>
<td></td>
</tr>
<tr>
<td>GSTP1</td>
<td>glutathione S-transferase pi</td>
<td></td>
</tr>
<tr>
<td>HIC1</td>
<td>hypermethylated in cancer 1</td>
<td>ZBTB29</td>
</tr>
<tr>
<td>HOXA9</td>
<td>homeobox A9</td>
<td></td>
</tr>
<tr>
<td>HOXB5</td>
<td>homeobox B5</td>
<td></td>
</tr>
<tr>
<td>HRAS</td>
<td>v-Ha-ras Harvey rat sarcoma viral oncogene homolog</td>
<td></td>
</tr>
<tr>
<td>HRK</td>
<td>harakiri, BCL2 interacting protein (contains only BH3 domain)</td>
<td>DP5</td>
</tr>
<tr>
<td>HS3ST2</td>
<td>heparan sulfate (glucosamine) 3-O-sulfotransferase 2</td>
<td>3OST2</td>
</tr>
<tr>
<td>HTLF</td>
<td>human T-cell leukemia virus enhancer factor</td>
<td>FOXN2</td>
</tr>
<tr>
<td>ID4</td>
<td>inhibitor of DNA binding 4, dominant negative helix-loop-helix protein</td>
<td></td>
</tr>
<tr>
<td>IGF2</td>
<td>insulin-like growth factor 2 (somatomedin A)</td>
<td></td>
</tr>
<tr>
<td>KRAS2</td>
<td>v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog</td>
<td>KRAS, KRAS1</td>
</tr>
<tr>
<td>LCN2</td>
<td>lipocalin 2 (oncogene 24p3)</td>
<td>NGAL</td>
</tr>
<tr>
<td>LDHA</td>
<td>lactate dehydrogenase A</td>
<td></td>
</tr>
<tr>
<td>LPHN2</td>
<td>latrophilin 2</td>
<td>KIAA0786, LEC1</td>
</tr>
<tr>
<td>MAGE</td>
<td>melanoma antigen family A</td>
<td></td>
</tr>
<tr>
<td>MBD1</td>
<td>methyl-CpG binding domain protein 1</td>
<td>PCMI, CXXC3</td>
</tr>
<tr>
<td>MBD2</td>
<td>methyl-CpG binding domain protein 2</td>
<td></td>
</tr>
<tr>
<td>MBD3</td>
<td>methyl-CpG binding domain protein 3</td>
<td></td>
</tr>
<tr>
<td>MECP2</td>
<td>methyl CpG binding protein 2</td>
<td></td>
</tr>
<tr>
<td>MEST</td>
<td>mesoderm specific transcript homolog (mouse)</td>
<td>PEG1</td>
</tr>
<tr>
<td>MGMT</td>
<td>O-6-methylguanine-DNA methyltransferase</td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>MLH1</td>
<td>mutL homolog 1, colon cancer, nonpolyposis type (E. coli)</td>
<td></td>
</tr>
<tr>
<td>MSH2</td>
<td>mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)</td>
<td></td>
</tr>
<tr>
<td>MSH3</td>
<td>mutS homolog 3 (E. coli)</td>
<td></td>
</tr>
<tr>
<td>MSH6</td>
<td>mutS homolog 6 (E. coli)</td>
<td></td>
</tr>
<tr>
<td>MSLN</td>
<td>mesothelin</td>
<td></td>
</tr>
<tr>
<td>MSX1</td>
<td>msh homeo box homolog 1 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>MSX2</td>
<td>msh homeo box homolog 2 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>MUTYH</td>
<td>mutY homolog (E. coli)</td>
<td></td>
</tr>
<tr>
<td>MYC</td>
<td>v-myc myelocytomatosis viral oncogene homolog c-MYC (avian)</td>
<td></td>
</tr>
<tr>
<td>MYOD1</td>
<td>myogenic factor 3</td>
<td></td>
</tr>
<tr>
<td>NDRG1</td>
<td>N-myc downstream regulated gene</td>
<td></td>
</tr>
<tr>
<td>NEDD1</td>
<td>neural precursor cell expressed, developmentally down-regulated 1</td>
<td></td>
</tr>
<tr>
<td>NEX3-1</td>
<td>NK3 transcription factor related, locus 1 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>NME1</td>
<td>non-metastatic cells 1, protein (NM23A) expressed in</td>
<td></td>
</tr>
<tr>
<td>NME2</td>
<td>non-metastatic cells 2, protein (NM23B) expressed in</td>
<td></td>
</tr>
<tr>
<td>NR3C1</td>
<td>nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)</td>
<td></td>
</tr>
<tr>
<td>PAGE</td>
<td>P antigen family</td>
<td></td>
</tr>
<tr>
<td>PGR</td>
<td>progesterone receptor</td>
<td></td>
</tr>
<tr>
<td>PIK3CG</td>
<td>phosphoinositide-3-kinase, catalytic, gamma polypeptide</td>
<td></td>
</tr>
<tr>
<td>PRDM2</td>
<td>PR domain containing 2, with ZNF domain</td>
<td></td>
</tr>
<tr>
<td>PSCA</td>
<td>prostate stem cell antigen</td>
<td></td>
</tr>
<tr>
<td>PTEN</td>
<td>phosphatase and tensin homolog (mutated in multiple advanced cancers 1)</td>
<td></td>
</tr>
<tr>
<td>PTGIS</td>
<td>prostaglandin I2 (prostacyclin) synthase</td>
<td></td>
</tr>
<tr>
<td>PTGS2</td>
<td>prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)</td>
<td></td>
</tr>
<tr>
<td>PTPRO</td>
<td>protein tyrosine phosphatase, receptor type, O</td>
<td></td>
</tr>
<tr>
<td>PYCARD</td>
<td>PYD and CARD domain containing</td>
<td></td>
</tr>
<tr>
<td>RAB32</td>
<td>RAB32, member RAS oncogene family</td>
<td></td>
</tr>
<tr>
<td>RARB</td>
<td>retinoic acid receptor, beta</td>
<td></td>
</tr>
<tr>
<td>RASSF1A</td>
<td>Ras association (RalGDS/AF-6) domain family 1, NORE2A, REH3P21, RDA32, A</td>
<td></td>
</tr>
<tr>
<td>RASSF2A</td>
<td>Ras association (RalGDS/AF-6) domain family 2, KIAA0168 A</td>
<td></td>
</tr>
<tr>
<td>RB1</td>
<td>retinoblastoma 1 (including osteosarcoma)</td>
<td></td>
</tr>
<tr>
<td>RBP1</td>
<td>retinol binding protein 1, cellular</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>RHOB</td>
<td>ras homolog gene family, member B</td>
<td></td>
</tr>
<tr>
<td>RND1</td>
<td>Rho family GTPase 1</td>
<td></td>
</tr>
<tr>
<td>ROBO1</td>
<td>roundabout, axon guidance receptor, homolog 1 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>RRAS</td>
<td>related RAS viral (r-ras) oncogene homolog</td>
<td></td>
</tr>
<tr>
<td>RUNX3</td>
<td>runt-related transcription factor 3</td>
<td></td>
</tr>
<tr>
<td>S100A4</td>
<td>S100 calcium binding protein A4</td>
<td></td>
</tr>
<tr>
<td>SCGB3A1</td>
<td>secretoglobin, family 3A, member 1</td>
<td></td>
</tr>
<tr>
<td>SERPINB5</td>
<td>serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5</td>
<td></td>
</tr>
<tr>
<td>SEZ6L</td>
<td>seizure related 6 homolog (mouse)-like</td>
<td></td>
</tr>
<tr>
<td>SFRP1</td>
<td>secreted frizzled-related protein 1</td>
<td></td>
</tr>
<tr>
<td>SFRP2</td>
<td>secreted frizzled-related protein 2</td>
<td></td>
</tr>
<tr>
<td>SFRP4</td>
<td>secreted frizzled-related protein 4</td>
<td></td>
</tr>
<tr>
<td>SFRP5</td>
<td>secreted frizzled-related protein 5</td>
<td></td>
</tr>
<tr>
<td>SLC5A8</td>
<td>solute carrier family 5 (iodide transporter), member 8</td>
<td></td>
</tr>
<tr>
<td>SLIT2</td>
<td>slit homolog 2 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>SMAD2</td>
<td>SMAD, mothers against DPP homolog 2 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>SMAD4</td>
<td>SMAD, mothers against DPP homolog 4 (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>SMARCC2</td>
<td>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2</td>
<td></td>
</tr>
<tr>
<td>SNCG</td>
<td>synuclein, gamma (breast cancer-specific protein 1)</td>
<td></td>
</tr>
<tr>
<td>SNRPN</td>
<td>small nuclear ribonucleoprotein polypeptide N</td>
<td></td>
</tr>
<tr>
<td>SOCS1</td>
<td>suppressor of cytokine signaling 1</td>
<td></td>
</tr>
<tr>
<td>SOCS3</td>
<td>suppressor of cytokine signaling 3</td>
<td></td>
</tr>
<tr>
<td>SORBS1</td>
<td>sorbin and SH3 domain containing 1</td>
<td></td>
</tr>
<tr>
<td>STK11</td>
<td>serine/threonine kinase 11 (Peutz-Jeghers syndrome)</td>
<td></td>
</tr>
<tr>
<td>TCF4</td>
<td>transcription factor 4</td>
<td></td>
</tr>
<tr>
<td>TFF2</td>
<td>trefoil factor 2 (spasmolytic protein 1)</td>
<td></td>
</tr>
<tr>
<td>TGFBR2</td>
<td>transforming growth factor, beta receptor II (70/80kDa)</td>
<td></td>
</tr>
<tr>
<td>THBS1</td>
<td>thrombospondin 1</td>
<td></td>
</tr>
<tr>
<td>TIMP3</td>
<td>tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory)</td>
<td></td>
</tr>
<tr>
<td>TMEFF2</td>
<td>transmembrane protein with EGF-like and two follistatin-like domains 2</td>
<td></td>
</tr>
<tr>
<td>TP53</td>
<td>tumor protein p53 (Li-Fraumeni syndrome)</td>
<td></td>
</tr>
<tr>
<td>TP73</td>
<td>tumor protein p73</td>
<td></td>
</tr>
<tr>
<td>TSIX</td>
<td>X (inactive)-specific transcript, antisense</td>
<td></td>
</tr>
<tr>
<td>TUSC3</td>
<td>tumor suppressor candidate 3</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name Description</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td><em>VHL</em></td>
<td>von Hippel-Lindau tumor suppressor VHL1</td>
<td></td>
</tr>
<tr>
<td><em>WIF1</em></td>
<td>WNT inhibitory factor 1</td>
<td></td>
</tr>
<tr>
<td><em>WT1</em></td>
<td>Wilms tumor WT</td>
<td></td>
</tr>
<tr>
<td><em>XAGE</em></td>
<td>X antigen family</td>
<td></td>
</tr>
<tr>
<td><em>XCE</em></td>
<td>X chromosome controlling element</td>
<td></td>
</tr>
<tr>
<td><em>XIST</em></td>
<td>X (inactive)-specific transcript DXS1089, swd66</td>
<td></td>
</tr>
<tr>
<td><em>XPA</em></td>
<td>xeroderma pigmentosum, complementation group XPAC, XP1 A</td>
<td></td>
</tr>
<tr>
<td><em>YWHAS</em></td>
<td>tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, sigma polypeptide 14-3-3-sigma</td>
<td></td>
</tr>
</tbody>
</table>

Gene symbols and gene names approved by the Human Gene Nomenclature Committee.
* p16^{INK4a} and p14^{ARF} are both transcribed from the CDKN2A locus. The approved gene symbol is CDKN2A for both these transcripts. In order to avoid confusion, the aliases p16^{INK4a} and p14^{ARF} are used throughout this thesis, except for paper I. In paper one, we have only analyzed p16^{INK4a} and not p14^{ARF}, the symbol CDKN2A therefore represents p16^{INK4a} in this paper.
Appendix III. Glossary

5-AZA-2'-DEOXYCYTIDINE A potent and specific inhibitor of DNA methylation.

BIOMARKER In cancer research and detection, a biomarker refers to a substance or process that is indicative of the presence of cancer in the body. It might be either a molecule secreted by a malignancy itself, or it can be a specific response of the body to the presence of cancer.

BISULPHITE SEQUENCING Allows a positive display of 5-methylcytosines in the gene promoter after sodium bisulphite modification as unmethylated cytosines appear as thymines, whereas 5-methylcytosines appear as cytosines in the final sequence.

CHROMATIN The complex of DNA and protein in the cell nucleus. See nucleosome.

CHROMATIN MODIFICATION Includes processes such as DNA methylation and histone modification (acetylation, phosphorylation, methylation and ubiquitylation).

CHROMATIN REMODELLING Transient changes in chromatin accessibility.

CHROMATIN-REMODELLING COMPLEX A polypeptide complex that can compact or relax the secondary and tertiary structure of chromatin.

CHROMOSOMAL INSTABILITY (CIN) Genetic instability in tumors resulting in high rates of chromosomal losses and gains.

CORE HISTONES These are histones H2A, H2B, H3 and H4. A nucleosome contains two copies of each of the core histones wrapped by 146-bp DNA.

CpG ISLAND First described by Adrian Bird as an unmethylated HpaII tiny fragment (HTF) and formerly commonly defined as a contiguous window of DNA of at least 200 base pairs in which the G:C content is at least 50% and the ratio of observed CpG frequency over the expected frequency exceeds 0.6. Recently, a more stringent definition of a 500-base-pair window with a G:C content of at least 55% and an observed over expected CpG frequency of at least 0.65 has been proposed to exclude most Alu repeat sequences.
DIFFERENTIALLY METHYLATED REGION (DMR) DNA segments in imprinted genes that show different methylation patterns between paternal and maternal alleles. Some DMRs acquire DNA methylation in the germ cells, whereas others acquire DNA methylation during embryogenesis.

DNA HYPERMETHYLATION An increased level of DNA methylation in a DNA sample at either an individual CpG dinucleotide or at a group of CpG dinucleotides relative to a reference DNA sample, usually derived from a normal tissue.

DNA HYPOMETHYLATION A decreased level of DNA methylation in a DNA sample at either an individual CpG dinucleotide or at a group of CpG dinucleotides (or even the entire genome) relative to a reference DNA sample, usually derived from a normal tissue.

EPIGENETIC Any heritable influence (in the progeny of cells or of individuals) on chromosome or gene function that is not accompanied by a change in DNA sequence. Examples of epigenetic events include mammalian X-chromosome inactivation, imprinting, centromere inactivation and position effect variegation.

EUCHROMATIN The lightly staining regions of the nucleus that generally contain decondensed, transcriptionally active regions of the genome.

GENOMIC IMPRINTING Process by which genes are selectively expressed by the maternal or paternal homologue of a chromosome.

HETEROCHROMATIN A cytologically defined genomic component that contains repetitive DNA (highly repetitive satellite DNA, transposable elements and ribosomal DNA gene clusters) and some protein-coding genes.

HETEROCHROMATIN PROTEIN 1 (HP1) A protein that binds to highly repetitive, heterochromatic satellite DNA at centromeres and telomeres.

HISTONES Small, highly conserved basic proteins, found in the chromatin of all eukaryotic cells, which associate with DNA to form a nucleosome.
**HOMEBOX** A 180-base-pair sequence that is present in many developmental genes. It encodes a DNA-binding helix–turn–helix motif, indicating that homeobox-containing gene products function as transcription factors.

**ISOCHROMOSOME** An abnormal chromosome having a median centromere and two identical arms.

**LOSS OF HETEROZYGOSITY** (LOH) In cells that carry a mutated allele of a tumour-suppressor gene, the gene becomes fully inactivated when the cell loses a large part of the chromosome carrying the wild-type allele. Regions with a high frequency of LOH are believed to harbor tumor-suppressor genes.

**METHYLATION SPECIFIC-PCR** (MSP) Method to analyze the DNA methylation status of groups of CpG sites within a CpG island. The technique comprises two parts: (1) sodium bisulphite conversion of unmethylated cytosine's to uracil under conditions whereby methylated cytosines remains unchanged and (2) detection of the bisulphite induced sequence differences by PCR using specific primer sets for both unmethylated and methylated DNA.

**MICROSATELLITE INSTABILITY** (MSI) In diploid tumors, genetic instability that is due to a high mutation rate, primarily in short microsatellite repeats. Cancers with the MSI phenotype are associated with defects in DNA-mismatch-repair genes.

**MICROSATELLITE REPEATS** A class of repetitive DNA that is made up of repeats that are 2–8 nucleotides in length. Their mutation is used as a marker of defective mismatch repair.

**MISMATCH REPAIR** A genomic system that detects and repairs incorrectly paired nucleotides that are introduced during DNA replication.

**NUCLEOSOME** The fundamental unit into which DNA and histones are packaged in eukaryotic cells. It is the basic structural subunit of chromatin and consists of 200 bp of DNA and an octamer of histone proteins, comprising two of each core histone.
PLURIPOTENCY The ability of a cell to contribute to several tissues in a developing organism. If a cell is able to contribute to all tissues, it is said to be totipotent.

PROMOTER HYPERMETHYLATION The addition of a methyl group to the cytosine ring at the 5th carbon position to form methyl cytosine throughout the CpG island of a gene promoter. The cytosines that become methylated are 5' to guanine.

SODIUM BISULPHITE TREATMENT OF DNA Leads to sequence variations as unmethylated but not methylated cytosines are converted to uracil.

X-INACTIVATION Mammalian females have two X chromosomes per genome, whereas males have only one. In female mammals, one X chromosome is functionally silenced during embryogenesis to ensure that the stoichiometry of X-chromosomal and autosomal gene products is the same in males and females.

The definitions in this glossary are obtained from ref [1-6]

References for appendix III