High-throughput DNA methylation analysis in colorectal cancer and childhood leukemia

Eddy H. J. van Roon
High-throughput DNA methylation analysis in colorectal cancer and childhood leukemia

PhD thesis, Leiden University, June 20, 2012

ISBN: 978-94-6182-120-1

Printed by: Off Page
Cover design: E. H. J. van Roon and P. P. C. van Roon

High-throughput DNA methylation analysis in colorectal cancer and childhood leukemia

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus, prof. mr. P.F. van der Heijden
volgens besluit van het College voor Promoties
te verdedigen op woensdag 20 juni 2012
klokke 15:00

door

Eddy Herman Jasper van Roon
geboren te Alphen aan den Rijn
in 1979

PROMOTIECOMMISSIE:

Promotores:	Prof. dr. H. Morreau
	Prof. dr. G. J. van Ommen
Co-promotor:	Dr. J. M. Boer (LUMC / Erasmus MC, Rotterdam)
Overige leden:	Dr. R. P. Kuiper (Radboud University Medical Center, Nijmegen)
	Prof. dr. C. J. Cornelisse (LUMC / Roosevelt Academy, Middelburg)
	Dr. R. W. Stam (Erasmus MC, Rotterdam)

The studies presented in this thesis were performed at the Department of Human Genetics and the Department of Pathology of the Leiden University Medical Center (LUMC). The studies described in this thesis were partially supported by the prof. A.A.H. Kassenaar Foundation.

Financial support for the publication of this thesis has been provided by the J.E. Jurriaanse Stichting and MRC-Holland.

In science it often happens that scientists say, "You know that's a really good argument; my position is mistaken," and then they actually change their minds and you never hear that old view from them again. They really do it. It doesn't happen as often as it should, because scientists are human and change is sometimes painful. But it happens every day. I cannot recall the last time something like that happened in politics or religion.

~ **Carl Sagan, 1987**

Voor mijn moeder

CONTENTS

Chapter 1

General Introduction

GENERAL INTRODUCTION

Epigenetics

Epigenetics (*epi-* from the Greek word *επί* meaning "over" or "above") refers to heritable meiotic and mitotic changes in gene expression that occur without a change in the DNA sequence. The best understood mechanisms that account for this form of expression regulation are DNA methylation and covalent modifications of histones.

DNA methylation

DNA methylation is a covalent modification of the fifth carbon within the cytosine DNA base; the resulting base is often referred to as the 'fifth base' in the human genome (Figure 1). In adult mammalian somatic cells, this modification occurs only on the cytosine in a CpG dinucleotide pair. The CpG notation is used to distinguish the linear sequence of a cytosine preceding a guanine bound by a phosphate from the complementary base pairing between a cytosine and guanine residue (Figure 2). The methylation of these CpGs is facilitated by the DNA methyltransferases DNMT1, DNMT3A and DNMT3B[1-4]. DNMT1 resides at the replication fork and methylates CpG dinucleotides in the newly synthesized strand, making this enzyme essential for maintaining DNA methylation patterns in proliferating cells[5-8]. DNMT3A and DNMT3B are required for *de novo* methylation during embryonic development[5-7].

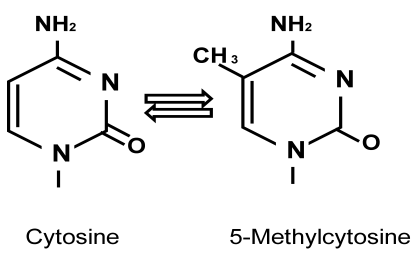


Figure 1 - Chemical structure of a cytosine nucleotide and 5-methylcytosine.

Due to spontaneous de-amination in the germ-line during evolution, CpG dinucleotides are rare within the genome[1]. However, CpG dinucleotides are enriched in DNA stretches ranging from 500 bp to several kb, and these regions are called CpG islands (GCIs)[1, 2, 4]. In contrast to the sparse CpG dinucleotides that occur throughout the genome, the majority of CGIs are hypomethylated. Approximately 60% of all genes contain a CGI within their promoter region that often expands to the first exon or intron and -regardless of the expression status of the associated geneare primarily unmethylated[4]. Although most CGIs reside in the 5' regions of genes, a large proportion of CGIs are located in inter-genic regions.

Hypermethylation of the promoter CGI is believed to down-regulate gene expression in two ways. First, DNA methylation may form a direct physical barrier against binding of the basic transcription complex or transcription enhancers (i.e., steric hindrance), thereby preventing downstream genes from being transcribed. Secondly, DNA methylation may

recruit methylation-specific proteins to the region, thus resulting in a cascade of silencing effects. Evidence for both hypotheses can be found in the literature[9]. CGI methylation is normally involved in allele-specific inactivation of imprinted genes and/or genes located on the inactive X chromosome, and aberrant CGI methylation has been found in numerous cancers[2, 10, 11].

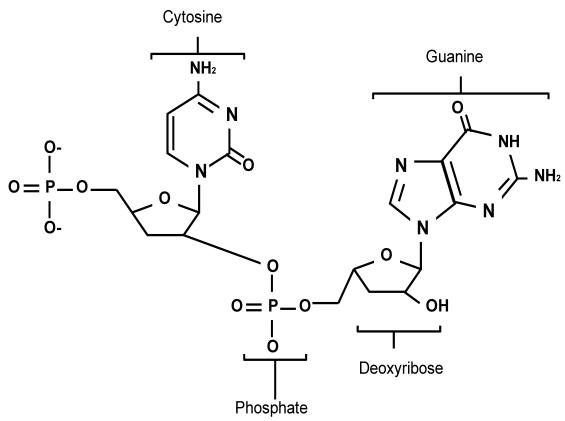


Figure 2 - Chemical structure of a CpG dinucleotide. The phosphate group (the p in CpG) indicates a deoxyribose bond between both nucleotides and thereby the 5'-3'locations of the cytosine and guanine. This annotation is used to prevent confusion with the hydrogen bonds between cytosine and guanine bases in complementary strands of DNA.

Histone modifications and chromatin state

In eukaryotes, genomic DNA is packaged with histone proteins into nucleosomes. A nucleosome consists of an octamer of histone proteins -comprised of two H2A-H2B heterodimers and two H3-H4 heterotetramers- that wrap ~146 bp of DNA around itself in 1.67 turns of a left-handed superhelix. Subsequently, these nucleosomes are themselves packed into chromatin, thus compacting DNA by approximately 10,000-fold. This 'packing' of two meters of DNA into a 1.7-µm cell nucleus is a considerable obstacle to replication, transcription and DNA repair complexes in reaching the DNA (Figure 3). To overcome this obstacle, dynamic changes in the chromatin state permit localized de-condensation from heterochromatin to euchromatin, thereby providing the nuclear machinery access to the DNA[12-16].

Condensed and de-condensed chromatin states coincide with a variety of post-translational covalent modifications of the core histone amino termini. A large number of histone modifications have been reported, among which acetylation, methylation, phosphorylation and -to a lesser extent- ubiquitination are the best characterized[12-17]. For all modifications (with the exception of arginine methylation), enzymes exist to either attach or remove the histone modification. An overview of histone modifications is presented in Table 1. The complexity of histone modifications -and our increasing understanding of their consequences- have led to the 'histone code' hypothesis. According to this hypothesis, histone modifications provide a platform for the binding of chromatin-associated regulators of gene expression[12-16].

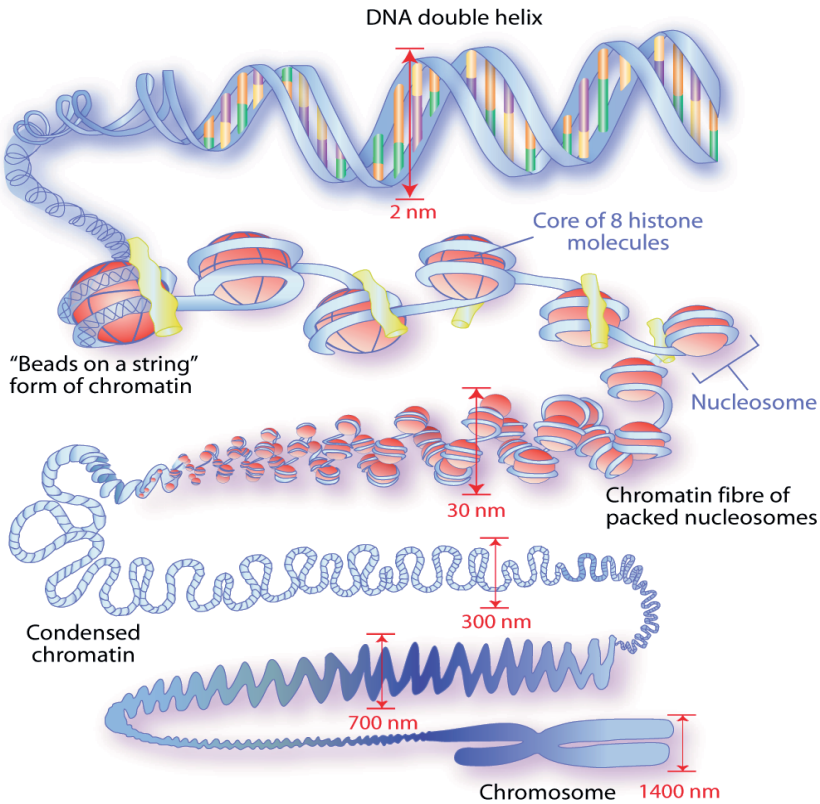


Figure 3 - Schematic representation of the sequential packaging of human DNA in the nucleus (adapted from www.epitron.eu)

Interaction between DNA methylation and histone modifications

Since epigenetic communication between DNA methylation and the chromatin state was initially described, the precise sequence of events that underlie this communication has been a subject of debate[18]. Currently, two progression models are considered to be plausible. The first model starts with initial DNA methylation that causes histone modifications via the recruitment of proteins that have methyl-DNA binding activity such as methyl-CpG-binding protein 2 (MeCP2), methyl-CpG-binding domain protein 1 (MDB1) and Kaiso (also known as the Zinc finger and BTB domain containing protein 33, or ZBTB 33). The subsequent recruitment of histone methyltransferases (HMTs) and histone deacetylases (HDACs) attach and detach histone modifications that are associated with transcriptional silencing and activation, respectively[19-25]. Finally, DNA methylation can inhibit active histone modification H3K4 methylation ($H3K4^{me}$)[26, 27].

Studies that support a model in which DNA methylation is initiated by histone modifications are increasing in number. These studies report that targets of the inactive histone modification $H3K27^{me3}$ and the enrichment of polycomb group 2 (PRC2) proteins in both embryonic (ES) and adult stem cells are pre-marked for *de novo* methylation in cancer[28-31]. Additional functional insights allowed the linking of PCR2 proteins, the presence of the

inactive histone mark H3K27^{me3} and absence of H3K4^{me3} to the recruitment of DNMTs and subsequent DNA methylation (Figure 4)[32-36]. The aforementioned studies led to a developmental model in which the balance between binding the mediators of inactivating histone mark H3K27^{me3}, PRC2 and the mediators of the activating histone mark H3K4^{me3}, the trithorax-group proteins, determine the DNA methylation and expression states of the regions to which they bind (Figure 4)[28-31, 37-40].

Although studies addressing this subject have not yielded conclusive evidence to support this model, they have revealed a high level of synergy between histone modifications and DNA methylation in regulating gene expression. Histone modifications are believed to act either sequentially or in combination with DNA methylation to generate the proposed histone code, which in turn conveys information to the nuclear machinery[15].

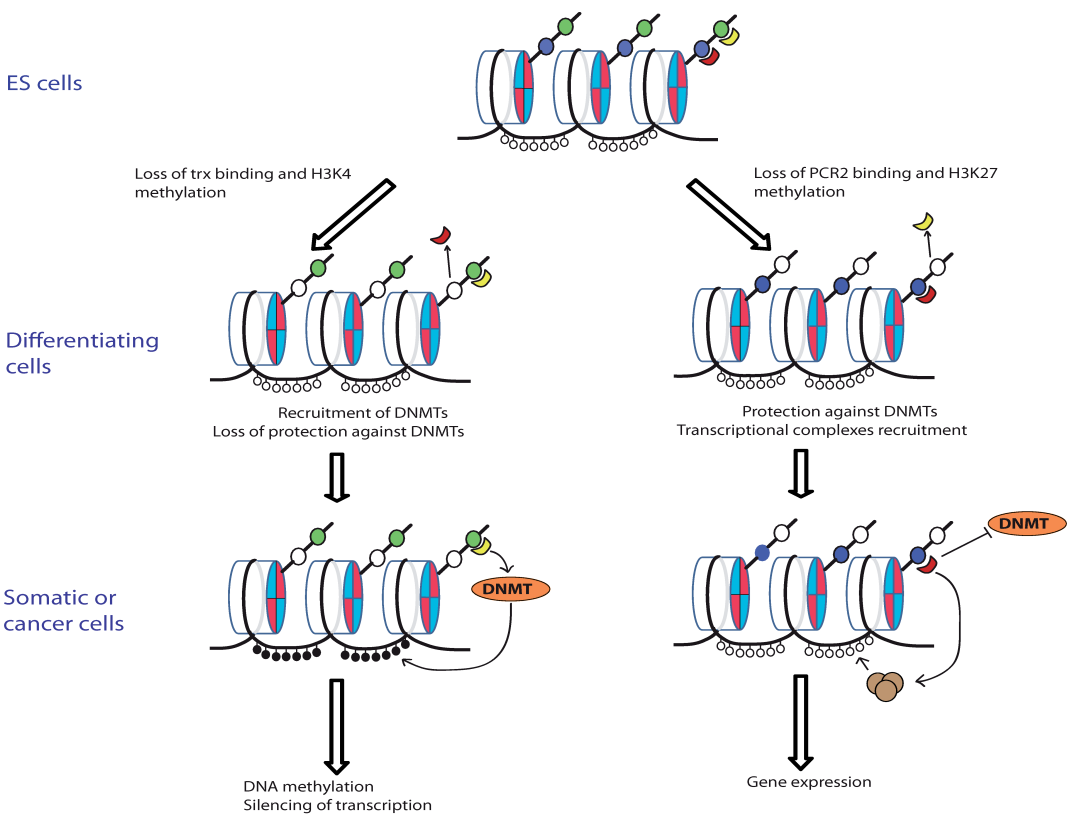


Figure 4 - Model of epigenetic regulation of gene expression in differentiation and tumorigenesis. Three nucleosomes that are composed of an H3-H4 hetero-tetramer (blue), two H2A-H2B dimers (red), the DNA (black line) with CpG dinucleotides (open circles attached to the DNA) and a histone tail with H3K4 (purple circle) and H3K27 (green circle) methylation are represented. A loss of PCR2 (yellow crescent) association during differentiation results in the loss of repressive H3K27 methylation, thereby allowing the binding of transcriptional complexes (light brown). The disassociation of trx family proteins (red) results in the loss of H3K4 methylation-mediated protection against DNMT (orange) recruitment. The remaining H3K27 methylation actively recruits the DNMT complexes, thereby resulting in methylation of the associated CpG dinucleotides (black circles attached to the DNA). The association of the trx or PCR2 complexes during differentiation can determine both the transcription of genes and downstream DNA methylation in somatic or cancer cells.

Table 1 - Histone modifications, locations and modifiers

Histone	Modification	Site	Enzyme	Proposed function
H2A	Acetylation	K5	TIP60/PLIP, HAT1, CBP/p300	Transcriptional activation
	Phosphorylation	S1	MSK1	Transcriptional repression
		T120	NHK-1	Mitosis
		S139	ATR, ATM, DNA-PK	DNA repair
	Ubiquitination	K119	HR6A	Spermatogenesis
H2B	Acetylation	K5	ATF2	Transcriptional activation
		K12	CBP/p300, ATF2	Transcriptional activation
		K15	CBP/p300, ATF2	Transcriptional activation
		K20	CBP/p300	Transcriptional activation
	Phosphorylation	S14	Mst1	Apoptosis
	Ubiquitination	K120	RNF20/hBRE1, RNF40, HR6A, HR6B,	Transcriptional activation
H3	Acetylation	K9	PCAF, GCN5	Transcriptional activation
		K14	PCAF, GCN5, TIP60/ PLIP, hTFIIIC90, TAF1, CBP/p300	Transcriptional activation
		K18	CBP/p300, PCAF, GCN5	Transcriptional activation
		K23	CBP/p300	Transcriptional activation
		K27	GCN5	Transcriptional activation
	Phosphorylation	T3	HASPIN	Mitosis
		S10	TG2, MSK1, MSK2	Transcriptional activation
		T11	DLK/ZIP	Mitosis
		S28	MSK1, MSK2	Transcription activation
	Methylation	K4	MLL($me^{1/2}$)	Transcriptional activation
			MLL2-4($me^{1/2/3}$)	Transcriptional activation
			SET1A, SET1B($me^{1/2/3}$)	Transcriptional activation
			SMYD3($me^{2/3}$)	Transcriptional activation
			SET7/9($me^{1/2}$)	Transcriptional activation
		K9	CLL8, RIZ1, SUV39h1, SYV39h2, ESET, G9A, EZH2	Transcriptional repression
		R17	CHARM1	Transcriptional activation
		K27	EZH2, G9A	Transcriptional silencing, X-inactivation (tri-methylation)
		K36	NSD1, SMYD2, SET2	Transcription activation, De-acetylation(single methylation)
		K79	DOT1L	Transcription activation, elongation / memory
H4	Acetylation	K5	HAT1, TIP60/PLIP, CBP/p300, HBO1	Transcriptional activation
		K8	TIP60/ PLIP, CBP/p300, HBO1	Transcriptional activation
		K12	HAT1, TIP60/PLIP, HBO1,	Transcriptional activation
		K16	TIP60/PLIP	
	Phosphorylation	S1	-	Mitosis
	Methylation	R3	PRMT1	Transcriptional activation
		K20	SET7/8, SUV4-20H1-2	Transcriptional repression

Chromatin state, activity and nuclear position

As mentioned above, chromatin status coincides with specific histone modifications (and thus to DNA methylation). These modifications are believed to regulate chromatin density either directly or by providing a surface substrate for interactions with other proteins[12, 16, 41]. Gene-rich and transcriptionally active regions can therefore be maintained as euchromatin, whereas gene-poor and transcriptionally inactive regions can be condensed to form heterochromatin.

Chromatin density -and thus transcriptional activity- is associated with specific interphase locations within the nucleus' volume. Heterochromatin generally clusters into condensed chromocenters that are located in the vicinity of the nucleolus, whereas active euchromatin is located in the central region and nuclear border[42]. This organization is not random, as differences have been reported based on cell type, shape, quiescence, commitment, functional status or transformation[43]. The availability of euchromatin to the interchromatin compartment -a channel network that is connected to the nuclear pores- has been postulated to facilitate transcription by the nuclear machinery that is located within this interchromatin compartment[44, 45]. Chromatin domains that contain transcriptionally active genes form euchromatic chromatin loops that migrate from the chromocenters to -or into- the interchromatin compartment[46-48].

Because histone modifications determine transcriptional activity and chromatin condensation, a reciprocal impact on nuclear architecture would be expected. Cremer et al. studied the relation between histone methylation and nuclear location in breast cancer interphase nuclei and reported clustering of histone methylation in close proximity to the nucleoli and -to a lesser extent- in the nuclear periphery[49]. Studies that investigated the relation between nuclear location and specific histone modifications for active (i.e., H3K4^{me3}, H4K20^{me1} and H4K20^{me3}) and inactive (i.e., H3K9^{me1}, H3K9^{me3} and H3K27^{me3}) chromatin revealed that methylation patterns are arranged in distinct nuclear layers, with a certain degree of overlap that depends on the type of epigenetic modification[50, 51].

Although the relations between gene activity, chromatic condensation and spatial location in the nucleus are less pronounced in quiescent cells than in proliferating cells, genomic loci that are found in the same chromosome territories during S phase are likely to be replicated at the same time and come into contact with the same chromatin factors following replication[52]. This provides a means to re-establish a given transcriptional and/or spatial pattern of organization in the daughter cells, as the factors that mediate the chromatin state are proposed to act coordinately on newly replicated loci[52]. As such, subnuclear compartments may not be critical for immediate biological events but may provide a mechanism for an accurate heritable transmission of the chromatin state and transcription patterns[52].

Lamina binding

A mechanism for anchoring chromatin to subnuclear compartments -and more specifically, to the nuclear envelope- occurs via binding of chromatin to the nuclear lamina (NL). The core of the NL consists of nucleus-specific, type V intermediate filament lamin proteins. These lamin proteins can be divided into A-type lamins, which are found predominately in differentiated cells, and B-type lamins, which are essential for cell viability[52, 53]. Stable

interactions between lamins and lamin-associated polypeptides (LAPs) are integral for both maintaining mechanical integrity of the nuclear envelope and providing anchor points for the aforementioned chromatin binding to the NL[53, 54]. The interaction between the NL and chromatin-associated proteins is mediated through LAPs, which bind to both the NL and to chromatin-associated proteins such as BAF, HP1 and Rb (see references 52, 54, and 55 for an overview).

Both genomic and proteomic experimental approaches have identified an association between the NL and heterochromatin. Although a putative role for the NL in the formation and/or maintenance of heterochromatin remains unclear, the NL is believed to anchor heterochromatin to the nuclear periphery, thereby providing structure and associated replication timing (Figure 5). A genetic approach to the study of B1-type lamin-associated DNA in human fibroblasts has identified 1,344 sharply defined DNA domains of 0.1-10 Mb each[56]. These lamina-associated domains (LADs) are characterized by hallmarks of heterochromatin such as a low level of gene expression, low gene density, high levels of H3K27^{me3} and low levels of H3K4^{me2}. Interestingly, these LADs are demarcated (Figure 5) by CpG islands, promoter regions driving transcription away from LADs and binding regions of the insulator protein CCCTC-binding factor (CTCF)[56].

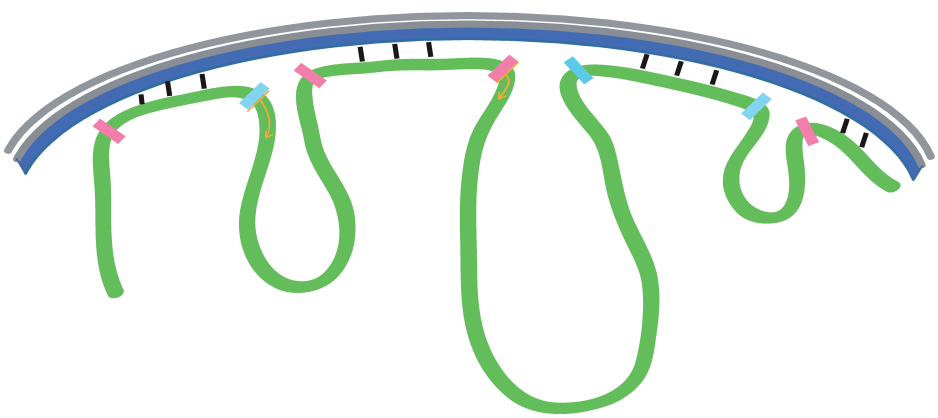


Figure 5 – Model of chromatin binding to the nuclear lamina. Large chromatin domains (green line) are dynamically associated (depicted as black lines) with the nuclear lamina (dark blue) adjacent to the nuclear envelope (gray). The LAD regions are demarcated by putative insulator elements, including CTCF binding sites (light blue), CpG islands (pink) and promoters that are orientated away from the lamina (orange arrows)[56]. Adapted from de Wit et al.[192].

The insulator protein CTCF

In vertebrates, CTCF is a ubiquitously expressed, 11-zinc finger protein that has been shown to bind to a larger number of binding sites in the genome; the number of binding sites ranges from 13,804 to 26,814 sites, depending on the cell type, technique and method of analysis[57-61]. This 'Jack-of-all–trades' protein has been implicated in diverse roles in gene regulation, including promoter activation/repression, enhancer blocking and/or barrier insulation, hormone-responsive silencing, genomic imprinting and -most recently- long-range chromatin interactions[62]. In addition to the aforementioned correlation between LAD boundaries and CTCF, a recent genome-wide mapping study uncovered a significant proportion of CTCF binding sites that are localized to the boundaries between euchromatic

and heterochromatic domains that are marked by H2AK5^{Ac} and H3K27^{me3}, respectively[61].

The discovery of CTCF-mediated intra- and inter-chromosome loop formation at the *IGF2/H19*[63, 64] and β-*globin* loci[65, 66] gives insight into how CTCF might form loops of condensed chromatin. Although the variability of CTCF loop formation by either homo- or hetero-dimerization with one of the many suggested protein partners makes it difficult to portray CTCF in a universal model, the high number and high variation of CTCF binding sites throughout the genome suggest a key role for CTCF in nuclear architecture. It has been reported recently that CTCF binding sites are generally located in chromatin linker regions that are flanked by at least 20 symmetrically distributed nucleosomes, thus revealing both a genome-wide role for CTCF in nucleosome positioning and a link to the regulation of chromatin structure[67]. Among CTCF's many protein partners, the recruitment of the Polycomb Repressor Complex 2 member Suz12 by DNA-bound CTCF is associated with the subsequent acquisition of H3K27^{me3}, indicating that CTCF binding might initiate local heterochromatin formation[68].

Studies of CTCF binding to the imprinting control region of *IGF2/H19* have shown that CTCF binding is DNA methylation sensitive[69, 70]. Additionally, methylation of a single CpG dinucleotide within the CTCF consensus sequence of the chicken *β-globin* gene is sufficient to block CTCF binding. This finding has led to the classification of CTCF binding sites into the following three groups: sites without CpG dinucleotides, sites that contain DNA methylation and unmethylated sites. A small-scale comparison between pre-B and thymocyte cell lines found that sites with unchanged CTCF occupancy are generally unmethylated, whereas sites that display differential binding between lineages may acquire CpG methylation[69, 71]. Not only does the binding of CTCF appear to be DNA methylation sensitive, but the recruitment and activation of the DNMT1 inhibitor PARP-1 by DNA-bound CTCF seem to indicate a protective function against methylation of CTCF binding sites that contain CpG dinucleotides[72, 73]. Interestingly, a specific subset of CTCF remains associated with chromosomes during mitosis, suggesting a possible role in the maintenance of epigenetic marks throughout cell division[74, 75]. Together with its insulator function, the protection of CTCF's own binding sites throughout cell division could link epigenetic transcriptional regulation and nuclear architecture and could explain epigenetic heritability through cell division in differentiated cells. Naturally occurring DNA sequence variations can also influence CTCF binding. For example, a polymorphism in a CTCF binding site downstream of *MMP-7* that leads to differential CTCF binding is a possible genetic factor in breast cancer[76].

DNA methylation in cancer

Aberrant methylation of CpG dinucleotides is commonly seen in cancer and -shown by studies of this phenomenon- is recognized as an important step in tumorigenesis[4, 77]. In carcinomas, hypomethylation of the genome is accompanied by regional hypermethylation of CGIs compared to the normal epithelium cells from which they arise[2, 4, 77]. Global hypomethylation has been linked to both genomic instability and increasing mutation rates, whereas hypermethylation of promoter CGIs can lead to transcriptional inactivation of the associated gene[78, 79]. This aberrant CGI hypermethylation is accompanied by the recruitment of methyl-CpG binding domain (MBD) proteins and histone deacetylases (HDACs) and is associated with histone modifications that are associated with expressional

down-regulation[80]. In various types of cancers, promoter hypermethylation of tumor suppressor genes (TSGs) such as *p16INK4a*[81-83], *MLH1*[84-87], *BRCA1*[88, 89] and *Rb*[90] have been described.

Hypermethylation of CGIs in tumors is part of a cascade that can lead to the down-regulation of expression through changes in the histone code and possibly even via the nuclear location of the associated DNA. Due to the robust nature of DNA methylation, changes in the DNA methylome can be detected using various techniques, and there exists a huge potential for the use of DNA methylation as a diagnostic and/or prognostic marker[91]. Additionally, the identification of aberrancies in epigenetic regulation might provide new insights into tumorigenesis and perhaps pave the way for the development and application of new cancer treatments that reverse DNA methylation.

The initiation of cancer-related DNA methylation has been a focus for researchers since it was first discovered. The aforementioned complex interplay between DNA methylation with histone modifications and their mediators yields a large group of epigenetic machinery proteins that can play a role in epigenetic tumorigenesis. A complete understanding of the initiation and impact of DNA methylation in tumorigenesis is needed to distinguish between randomly accumulated DNA methylation and the methylation of targets that are important in the development of cancer.

Colorectal cancer: clinical context

Colorectal cancer (CRC) is the third and second most common type of cancer in males and females, respectively, and one of the leading causes of cancer-related deaths in both Europe and the US[92, 93]. In the Netherlands, the lifetime risk for developing CRC is 6% (an incidence of approximately one in 17) among both genders. In recent years, the number of new CRC cases and associated deaths has seemingly decreased in developed countries, and this is possibly due to improved screening methods and early diagnosis[92, 93]. However, in Japan and other developing countries, the incidence of CRC is increasing, and this is believed to reflect a combination of factors that are related to a Western lifestyle, including changes in dietary patterns, obesity and an increased prevalence of smoking[92-96]. Worldwide, it is estimated that approximately one million new cases are diagnosed annually[92, 93, 96]. Over 95% of colorectal cancers are adenocarcinomas, and approximately half of these patients develop a local recurrence or a distant metastasis during the course of the disease. Survival depends greatly on early detection, particularly before the tumor has metastasized[97]. The five-year survival rate ranges from 93.2 to 82.5% for the early stages in which no lymph node metastasis has occurred yet[98]. In cases of lymph node metastasis (stage III; see www.UICC.org) or distant metastasis (stage IV), the survival rates are 59.5 and 8.1%, respectively. Stage III and stage IV tumors are typically treated with chemotherapy consisting of 5-fluorouracil compounds either with or without oxaliplatin or irinotecan[97, 99]. In recent years, insights into the molecular pathogenesis of colorectal cancer have led to the use of targeted therapeutics that are specific for the epidermal growth factor receptor (EGFR) and vascular endothelial growth factor (VEGF)[97, 99]. Although the success of these therapies in CRC is limited, these examples illustrate how molecular biological research contributes to the development of promising new therapies.

Tumorigenesis of CRC

The accumulation of genetic and epigenetic changes results in the progressive transformation of normal colon epithelium to hyperplasia, dysplasia and eventually adenocarcinoma. This stepwise progression of tumorigenesis in colorectal cancer has served as an example of other types of tumors. The recently updated yet classic Vogelgram[100] shows that colorectal neoplasias can be characterized based on molecular features. The predilection for specific molecular alterations at different sites in the colon is remarkable. Right-sided (proximal) and left-sided (distal) CRC[100-103] can be seen grossly as the following two classic and distinct genetic pathways (Figure 6): tumors with high levels of chromosomal instability (CIN) or microsatellite instability (MSI or MSI high/MSI-H). The CIN pathway (which comprises 50-70% of sporadic colon cancers) is characterized by a change in chromosomal copy number such as a chromosomal gain, loss or a copy-neutral loss of heterozygosity (cnLOH)[104]. Tumors that arise via this pathway are often located in the left-sided colon (i.e., distal to the splenic flexure) and are often aneuploid. Although these CIN colon tumors progress through the adenoma-carcinoma progression pathway, the facilitating mechanism is not completely understood. Specific mutations in genes that are involved in mitotic spindle checkpoints and DNA replication checkpoints (e.g., *hBUB1* and *hBUBR1*) have been proposed to underlie CIN, and self-propagating genomic instability can occur in the absence of genetic mutations[104-108]. To date, no data have been provided compelling evidence that mutations in any of these genes provide more than a permissive role for CIN, despite the tight association between CIN and mutant APC and p53[106].

Tumors that arise via the MSI pathway (comprising ~15% of sporadic colon cancers) are typically diploid, right-sided (i.e., before the splenic flexure) and carry small deletions and/or insertions in short repetitive sequences (A_n or CA_n, where n is the number of repeats) as a result of a loss of function of any of the DNA mismatch repair (MMR) genes[106]. In colon cancer, MSI is found in the context of Lynch syndrome (previously known as hereditary non-polyposis colorectal cancer, or HNPCC) with germline mutations in one of four MMR genes, primarily in *MLH1* or *MSH2*[109] and -to a lesser extent- in *MSH6*[110] or *PMS2*[111]. Deletions in *EPCAM/TACSTD1*, which is upstream of *MSH2*, cause sequential *MSH2* methylation[112, 113]. Although rare, several studies have described inherited and *de novo* germline methylation of *MLH1* in patients with Lynch-like colon cancer[114-120]. Approximately 15% of all sporadic colon cancers are due to somatic biallelic or hemiallelic methylation of the *MLH1* promoter[121].

A growing understanding of the impact and level of promoter, inter- and intra-gene CGI methylation that is described as aberrantly methylated in MSI colon cancer has led to the classification of colon cancers into the following CpG island methylator phenotypes (CIMP), regardless of MSI status: CIMP1 (CIMP-high), CIMP2 (CIMP-low) and CIMP0 (CIMP-negative)[4, 122-124]. Although the definition of CIMP has been debated in the literature, an integrated genetic and epigenetic analysis provided definitions for each of these three phenotypes[124]. The phenotype with the highest frequency of aberrant methylation, CIMP1, is associated with sporadic MSI, somatic *BRAF* mutations and the methylation of a debated set of methylation markers. The methylation status of the second phenotype, CIMP2, has also been the subject of debate. Methylation has been found among cancers in this group, albeit to a lesser extent than among CIMP1 tumors. Although methylation

markers have been suggested for both groups, indecisiveness regarding a defined marker set has led to *MLH1* methylation (and thereby sporadic MSI) and *BRAF* mutations as being the best indicators for CIMP1, whereas *KRAS* and *TP53* mutations are often found in CIMP2 and CIMP0 tumors, respectively[122-126].

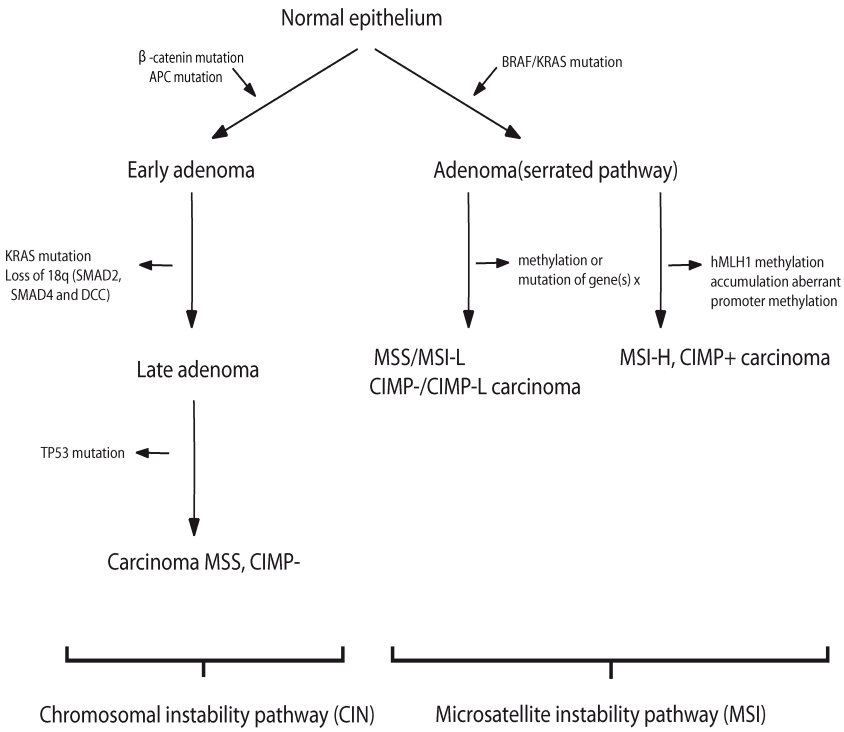


Figure 6 – A model of the CIN and MSI tumorigenesis pathways

The cause of aberrant DNA methylation in CRC

The underlying causes of aberrant methylation and subsequent sporadic MSI colon cancer remain largely unknown. Both *BRAF* and *KRAS* mutations have been observed in the earliest identified colonic neoplasms, and recent studies have provided evidence that induction of the ras oncogenic pathway results in DNA hypermethylation[127-132]. Although activating *KRAS* and *BRAF* mutations are present in early colonic neoplasia, they give rise to different types of polyps. *KRAS* mutations are primarily found in adenomatous polyps, whereas *BRAF* mutations occur primarily in polyps that have a serrated architecture and have been suggested as precursor lesions for MSI carcinomas[129, 130, 132-135]. In early neoplasia, *BRAF* mutation was are associated with CIMP, which has been suggested to precede MSI by *MLH1* promoter methylation[128-130, 132, 136]. This association of *BRAF* mutations with sporadic MSI colon cancer, their precursor lesions and CIMP (in contrast to *KRAS* mutations) suggests that the two mutations (*BRAF* and *KRAS*) follow distinct tumorigenesis pathways despite being members of the same signaling pathway[128, 130, 132, 136].

Although *KRAS* and *BRAF* mutations are observed in early colonic neoplasia, the sequence of events regarding DNA methylation remains unclear. Promoter methylation

of O6-methylguanine DNA methyltransferase (*MGMT*) often occurs in many tumor types, including colon cancer[137-139]. Additionally, epigenetic down-regulation of *MGMT* expression is often seen in tumor-adjacent normal colon mucosa[140]. MGMT is a DNA base excision repair protein that removes mutagenic and cytotoxic adducts from the O6 position of guanine. O6-methylguanide often mispairs with thymine during replication, resulting in the conversion from a GC pair to an AT pair if the adduct is not removed. Inactivation of the *MGMT* gene via promoter hypermethylation can result in G-to-A transitions in the mutational hotspots within codons 12 and 13 of the *KRAS* oncogene, as well as in *TP53*[137, 139, 140]. Therefore, methylation of the *MGMT* promoter might initiate tumor progression through secondary *KRAS* and/or *TP53* mutations, a theory that might argue against the initiation of aberrant DNA methylation via the occurrence of activating *KRAS* mutations.
Although *BRAF* mutations cannot be explained by *MGMT* inactivation, methylation of the *IGFBP7* promoter has been shown to facilitate the oncogenic potency of activated BRAF. Active IGFBP7 is required for oncogene-induced cellular senescence (OIS), an important tumor suppressor mechanism[141-143]. Escaping the OIS pathway could favor selection for activating *BRAF* mutations. The accumulation of aberrant promoter hypermethylation might provide a favorable environment for the oncogenicity of mutated *BRAF*, which could explain the association between *BRAF* mutations and CIMP. However, the association between *BRAF* mutations and MSI remains a molecular puzzle. More research is needed to determine the initiating factor and the role of *MLH1* methylation in this model.

MLL-rearranged B-lineage leukemia

Acute lymphoblastic leukemia (ALL) is the most common malignancy in children under the age of 15 and accounts for 26.8% of all childhood cancers[144, 145]. This lymphoid leukemia can be divided into B and T cell leukemia depending on the cancer cell lineage. Over past few decades, treatment with a combination of chemotherapies has led to a considerable decrease in childhood cancer-related deaths and a 5-year survival rate that is currently between 78 and 83% in developed countries[144, 145].
However, upon age stratification of childhood ALL, a subgroup of infants who are younger than one year of age at diagnosis only attains a 5-year survival rate of approximately 50%[146, 147]. Although complete remission is achieved in most of these patients, a high relapse rate is the principal cause of this decrease in survival odds[146, 147]. Approximately 80% of infants with ALL carry chromosomal translocations that involve the mixed lineage leukemia (*MLL*) gene and typically exhibit an immature CD10-negative precursor B-lineage immunophenotype[146-148]. Within this infant ALL subgroup, the presence of *MLL* rearrangements and an age of younger than six months are described as the most important factors for predicting poor outcome[146, 147].
The most prevalent chromosomal translocations in infant ALL patients are t(4;11), t(11;19) and t(9;11), which fuse the N terminus of MLL to the C-terminal regions of AF4, ENL and AF9[146, 149]. Interestingly, these different translocations are characterized by distinct mRNA levels[150, 151] and DNA methylation patterns[152]. Genome-wide studies of DNA methylation levels as well as studies into the functions of MLL and fusion partner proteins have indicated that epigenetic changes play a major role in *MLL*-rearranged ALL and might be the driving force behind the expression differences between the translocation-stratified groups and control samples.

The normal function of MLL

The human *MLL* gene was discovered in the early 1990s by isolating the chromosomal breakpoints at chromosome 11q, cytoband 23[153-156]. A sequence comparison revealed three regions of sequence similarity with the *Drosophila melanogaster* gene *trithorax* (*trx*); thus, both are members of the trithorax group, an evolutionarily conserved family of proteins[157]. Similar to the function of trx in *Drosophila*, in mammals MLL acts as a transcriptional regulator of the class I homeodomain (*Hox*) genes and counters the repressive effects of the Polycomb group (PcG) proteins (Figure 4)[158-161]. The *Hox* genes, in turn, are transcription factors that direct cell fate during development. *MLL* is ubiquitously expressed both during development and in most adult tissues, including myeloid and lymphoid cells, and is required for definitive hematopoiesis[162-164]. In both *Mll*$^{-/-}$ mice and *trx*$^{-/-}$ flies, *Hox* gene expression is initiated correctly but deteriorates during embryogenesis, suggesting an essential role in maintaining expression patterns following initiation by other factors[157].

Identification of the different active domains of the large (3,968 amino acids) MLL protein has provided much insight into how MLL-mediated transcriptional regulation is facilitated (Figure 7). The MLL protein is cleaved by the protease taspase I into 320-kDa N-terminal and 180-kDa C-terminal fragments, both of which are core components of the MLL complex[165-168]. Two N-terminal domains -a region of three AT-hook domains and a region containing a CXXC zinc-finger domain- are believed to be involved in DNA binding[169-172]. The AT hook domain is a minor groove DNA binding motif that preferentially recognizes DNA that is distorted with bends or kinks, whereas the CXXC domain is the major determinant of subnuclear localization and target gene selection and recognizes and binds specifically to unmethylated CpG dinucleotides[173-175]. Although MLL can bind directly to DNA, MLL recruitment to chromatin can be mediated by DNA-binding protein partners such as menin (encoded by the *MEN1* gene)[176]. In addition to the CXXC, another domain targets MLL to sites that are associated with active chromatin. A central region between the third and fourth fingers contains three cysteine-rich plant homeodomain (PHD) zinc fingers and a fourth divergent PHD finger. This bromodomain has been shown to bind lysine-acetylated histone-derived peptides, thus suggesting preferential binding to acetylated histones by MLL[170, 172, 177-179].

Although the MLL protein has been associated with proteins that suppress gene expression, the recruitment of MLL to chromatin is most often associated with transcriptional activation. Both of the activating domains -namely, the transcription activation (TA) domain and the SET [Su(var)3-9, enhancer of zeste, and trithorax] domain- are located on the protein's C terminus[169-171]. The activating functions of both of these domains are mediated through epigenetics; the SET domain is directly responsible for methylating H3K4, and the TA domain recruits the histone acetyltransferases CREB-binding protein (CBP) and p300[180-183].

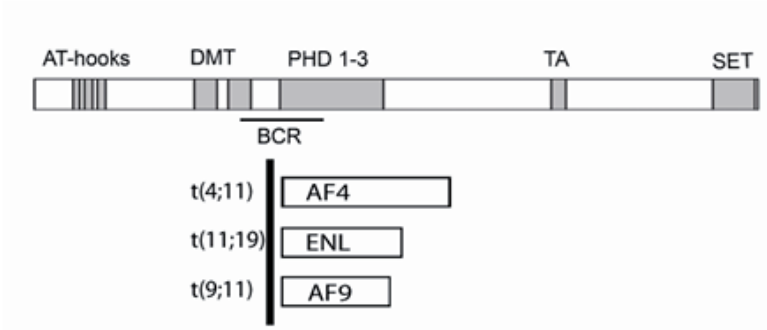


Figure 7 – Schematic representation of the MLL protein. The 89-kb *MLL* gene consists of 37 exons and encodes a 3,969-amino acid nuclear protein. MLL is cleaved at two cleavage sites (CS1 at amino acid 2666 and CS2 at amino acid 2718), resulting in two non-covalently associated subunits (N-terminal MLL (300 kDa) and C-terminal MLL (180 kDa)). The DNA-interacting domains (AT-hooks and the DNA methyltransferase homology domain (DMT) containing the zinc finger) are located in the N-terminal cleavage fragment. The PHD zinc-finger motifs facilitate the binding of proteins that are suggested to regulate MLL protein activity. This domain can be either present in an MLL fusion protein or completely absent, depending on the precise site of translocation in the breakpoint cluster region (BCR) spanning exons 8-13. Located on the C-terminal MLL domains are the transcriptional activation site (TA) and the SET domain (SET), both of which are involved in transferring marks of transcriptional activation to histone tails. The C-terminal parts of the fusion partners are shown beneath the MLL protein.

MLL fusion proteins: what do they add?

Given that MLL functions as an epigenetic transcriptional activator, a disruption in normal MLL function can be linked directly to the differences found in expression and DNA methylation between *MLL*-rearranged ALL groups and controls. All of the *MLL* rearrangements that have been found in ALL are believed to arise from a failure of DNA double-strand break repair during hematopoiesis. Most of the *MLL* rearrangements target the breakpoint cluster region that is located between exons 8 and 13, resulting in a fusion protein that contains N-terminal MLL and the C terminus of a fusion partner[184, 185]. Mouse studies have revealed that a truncation of MLL after exon 8 is not sufficient to induce leukemia but requires a functional C-terminal portion of a fusion protein[186]. The perturbed H3K4 methylation of one *MLL* copy is therefore not sufficient to initiate leukemogenesis. The aforementioned translocations with fusion partners *AF4*, *AF9* or *ENL* account for more than 80% of all *MLL*-rearranged leukemias, and all three resulting fusion partners contain a C-terminal transcriptional activation domain. These activation domains are associated with the H3K79 histone methyltransferase DOT1L[187-191]. H3K79 methylation levels are increased in targets that are crucial for *MLL*-rearranged leukemogenesis[187-191]. Given that the various methylation marks regulate transcription in unique ways[12, 14-16, 41], the addition of H3K79 methylation to normal H3K4 methylation at MLL-associated promoters could account for the aberrant over-expression and DNA methylation differences in *MLL*-rearranged leukemia. However, the high levels of DNA hypermethylation observed in patients with *MLL-AF4* and *MLL-ENL* translocations cannot currently be explained with the current understanding of *MLL* rearrangements in infant B-ALL.

SCOPE AND OUTLINE OF THIS THESIS

Tumor formation is the result of either DNA mutations in the genetic code, of chromosomal alterations or of epigenetic changes, the latter with DNA hyper- and hypomethylation. DNA mutations comprise base substitutions (point mutations) as well as relatively small insertions and deletions. Chromosomal alterations can occur as copy number variations, translocations and inversion of chromosomes. Often these alterations occur in parallel to ploidy changes of the whole genome of the cells.

Epigenetics, comprising DNA and histone modifications, is a relatively new field of study and recent technical possibilities have fuelled a growing interest in the role of epigenetics in tumorigenesis. Although the causes and effects of DNA hyper- and hypo-methylation in cancer are still being investigated, cancer-specific methylation profiles can be potentially used for clinical purposes such as pre-symptomatic screening for colorectal cancer in serum and faeces. In this thesis DNA methylation was studied in colon cancer and infant acute lymphoblastic leukemia.

In **Chapter 1** an introduction is given on this topic. In **Chapter 2**, using the differential methylation hybridization (DMH) technique, home-spotted CpG island microarrays were employed on right-sided colon cancer samples and compared with normal colon mucosa. High frequent methylation of the *PTPRGint1* sequence different types of colon cancer was seen. The *PTPRGint1* sequence turned out to be a binding site for CTCF, a protein that is involved in regulation of chromatin modifications. In **Chapter 3** a relatively young cohort of colon cancer patients with *MLH1* promoter hypermethylation was studied. Interestingly, this epigenetic down-regulation of *MLH1* is mostly seen in elderly colon cancer patients above 70 years of age. In **Chapter 4** we used the DMH technique in combination with a high density oligonucleotide CpG island microarray to obtain methylation profiles of colon cancer samples and matching normal colonic mucosa. As DNA methylation is suggested to be a consequence of pre-existing histone modifications we filtered *BRAF* mutation-specific methylation profiles for such pre-marking and identified promoter methylation of *FOX* genes involved in oncogene induced senescence.

Finally, the CpG Island microarrays were employed to study cancer-specific DNA methylation in *MLL*-rearranged B-ALL **(Chapter 5).** Infant ALL-specific as well as *MLL*-translocation-specific promoter methylation patterns were identified. These promoter methylation patterns correlated strongly with expression and outcome.

References

1. Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986;321:209-213.
2. Illingworth RS, Bird AP. CpG islands--'a rough guide'. FEBS Lett 2009;583:1713-1720.
3. Robert MF, Morin S, Beaulieu N, Gauthier F, Chute IC, Barsalou A, MacLeod AR. DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. Nat Genet 2003;33:61-65.
4. Toyota M, Issa JP. CpG island methylator phenotypes in aging and cancer. Semin Cancer Biol 1999;9:349-357.
5. Bestor TH. The DNA methyltransferases of mammals. Hum Mol Genet 2000;9:2395-2402.
6. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007;128:683-692.
7. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:247-257.
8. Surani MA. Reprogramming of genome function through epigenetic inheritance. Nature 2001;414:122-128.
9. Bogdanovic O, Veenstra GJ. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma 2009;118:549-565.
10. Barlow DP. Gametic imprinting in mammals. Science 1995;270:1610-1613.
11. Singer-Sam J, Riggs AD. X chromosome inactivation and DNA methylation. EXS 1993;64:358-384.
12. Berger SL. Histone modifications in transcriptional regulation. Curr Opin Genet Dev 2002;12:142-148.
13. Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol 2007;14:1008-1016.
14. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell 2007;130:77-88.
15. Jenuwein T, Allis CD. Translating the histone code. Science 2001;293:1074-1080.
16. Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem 2006;75:243-269.
17. Peterson CL, Laniel MA. Histones and histone modifications. Curr Biol 2004;14:R546-R551.
18. Razin A, Cedar H. Distribution of 5-methylcytosine in chromatin. Proc Natl Acad Sci U S A 1977;74:2725-2728.
19. Sasai N, Nakao M, Defossez PA. Sequence-specific recognition of methylated DNA by human zinc-finger proteins. Nucleic Acids Res 2010.
20. Sansom OJ, Maddison K, Clarke AR. Mechanisms of disease: methyl-binding domain proteins as potential therapeutic targets in cancer. Nat Clin Pract Oncol 2007;4:305-315.
21. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 2006;31:89-97.
22. Ballestar E, Esteller M. Methyl-CpG-binding proteins in cancer: blaming the DNA methylation messenger. Biochem Cell Biol 2005;83:374-384.
23. Jorgensen HF, Bird A. MeCP2 and other methyl-CpG binding proteins. Ment Retard Dev Disabil Res Rev 2002;8:87-93.
24. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 1998;393:386-389.
25. Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 1998;18:6538-6547.
26. Okitsu CY, Hsieh CL. DNA methylation dictates histone H3K4 methylation. Mol Cell Biol 2007;27:2746-2757.

27. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457-466.
28. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007;39:237-242.
29. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. Epigenetic stem cell signature in cancer. Nat Genet 2007;39:157-158.
30. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 2007;39:232-236.
31. McCabe MT, Lee EK, Vertino PM. A multifactorial signature of DNA sequence and polycomb binding predicts aberrant CpG island methylation. Cancer Res 2009;69:282-291.
32. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res 2003;31:2305-2312.
33. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 2003;13:1192-1200.
34. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van EA, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di CL, de LY, Fuks F. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006;439:871-874.
35. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448:714-717.
36. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO Rep 2009;10:1235-1241.
37. Cloos PA, Christensen J, Agger K, Helin K. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev 2008;22:1115-1140.
38. Issaeva I, Zonis Y, Rozovskaia T, Orlovsky K, Croce CM, Nakamura T, Mazo A, Eisenbach L, Canaani E. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Mol Cell Biol 2007;27:1889-1903.
39. Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, Chen S, Iwase S, Alpatov R, Issaeva I, Canaani E, Roberts TM, Chang HY, Shi Y. A histone H3 lysine 27 demethylase regulates animal posterior development. Nature 2007;449:689-694.
40. Rada-Iglesias A, Enroth S, Andersson R, Wanders A, Pahlman L, Komorowski J, Wadelius C. Histone H3 lysine 27 trimethylation in adult differentiated colon associated to cancer DNA hypermethylation. Epigenetics 2009;4:107-113.
41. Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol 2007;14:1008-1016.
42. Sadoni N, Langer S, Fauth C, Bernardi G, Cremer T, Turner BM, Zink D. Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. J Cell Biol 1999;146:1211-1226.
43. Folle GA. Nuclear architecture, chromosome domains and genetic damage. Mutat Res 2008;658:172-183.

44. Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B, Schrock E, Speicher MR, Mathieu U, Jauch A, Emmerich P, Scherthan H, Ried T, Cremer C, Lichter P. Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol 1993;58:777-792.
45. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2001;2:292-301.
46. Bartova E, Krejci J, Harnicarova A, Kozubek S. Differentiation of human embryonic stem cells induces condensation of chromosome territories and formation of heterochromatin protein 1 foci. Differentiation 2008;76:24-32.
47. Chambeyron S, Bickmore WA. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes Dev 2004;18:1119-1130.
48. Wiblin AE, Cui W, Clark AJ, Bickmore WA. Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. J Cell Sci 2005;118:3861-3868.
49. Cremer M, Zinner R, Stein S, Albiez H, Wagler B, Cremer C, Cremer T. Three dimensional analysis of histone methylation patterns in normal and tumor cell nuclei. Eur J Histochem 2004;48:15-28.
50. Skalnikova M, Bartova E, Ulman V, Matula P, Svoboda D, Harnicarova A, Kozubek M, Kozubek S. Distinct patterns of histone methylation and acetylation in human interphase nuclei. Physiol Res 2007;56:797-806.
51. Zinner R, Albiez H, Walter J, Peters AH, Cremer T, Cremer M. Histone lysine methylation patterns in human cell types are arranged in distinct three-dimensional nuclear zones. Histochem Cell Biol 2006;125:3-19.
52. Taddei A, Hediger F, Neumann FR, Gasser SM. The function of nuclear architecture: a genetic approach. Annu Rev Genet 2004;38:305-345.
53. Stuurman N, Heins S, Aebi U. Nuclear lamins: their structure, assembly, and interactions. J Struct Biol 1998;122:42-66.
54. Schirmer EC, Foisner R. Proteins that associate with lamins: many faces, many functions. Exp Cell Res 2007;313:2167-2179.
55. Schardin M, Cremer T, Hager HD, Lang M. Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. Hum Genet 1985;71:281-287.
56. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de KA, Wessels L, de LW, van SB. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 2008;453:948-951.
57. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell 2007;129:823-837.
58. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K. Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. Nucleic Acids Res 2008;36:5221-5231.
59. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 2007;128:1231-1245.
60. Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet 2001;17:520-527.
61. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res 2009;19:24-32.
62. Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell 2009;137:1194-1211.

63. Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. Proc Natl Acad Sci U S A 2006;103:10684-10689.
64. Lewis A, Murrell A. Genomic imprinting: CTCF protects the boundaries. Curr Biol 2004;14:R284-R286.
65. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de LW. The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 2003;35:190-194.
66. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de LW. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006;20:2349-2354.
67. Fu Y, Sinha M, Peterson CL, Weng Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. PLoS Genet 2008;4:e1000138.
68. Han L, Lee DH, Szabo PE. CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. Mol Cell Biol 2008;28:1124-1135.
69. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 2000;405:482-485.
70. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 2000;405:486-489.
71. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 2008;132:422-433.
72. Caiafa P, Zlatanova J. CCCTC-binding factor meets poly(ADP-ribose) polymerase-1. J Cell Physiol 2009;219:265-270.
73. Guastafierro T, Cecchinelli B, Zampieri M, Reale A, Riggio G, Sthandier O, Zupi G, Calabrese L, Caiafa P. CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. J Biol Chem 2008;283:21873-21880.
74. Burke LJ, Zhang R, Bartkuhn M, Tiwari VK, Tavoosidana G, Kurukuti S, Weth C, Leers J, Galjart N, Ohlsson R, Renkawitz R. CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. EMBO J 2005;24:3291-3300.
75. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A 2008;105:8309-8314.
76. Beeghly-Fadiel A, Long JR, Gao YT, Li C, Qu S, Cai Q, Zheng Y, Ruan ZX, Levy SE, Deming SL, Snoddy JR, Shu XO, Lu W, Zheng W. Common MMP-7 polymorphisms and breast cancer susceptibility: a multistage study of association and functionality. Cancer Res 2008;68:6453-6459.
77. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis 2010;31:27-36.
78. Baylin SB, Herman JG. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. Trends Genet 2000;16:168-174.
79. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999;21:163-167.
80. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet 2007;8:286-298.
81. Herman JG, Merlo A, Mao L, Lapidus RG, Issa JP, Davidson NE, Sidransky D, Baylin SB. Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. Cancer Res 1995;55:4525-4530.

82. Little M, Wainwright B. Methylation and p16: suppressing the suppressor. Nat Med 1995;1:633-634.
83. Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, Burger PC, Baylin SB, Sidransky D. 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nat Med 1995;1:686-692.
84. Cunningham JM, Christensen ER, Tester DJ, Kim CY, Roche PC, Burgart LJ, Thibodeau SN. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res 1998;58:3455-3460.
85. Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JP, Markowitz S, Willson JK, Hamilton SR, Kinzler KW, Kane MF, Kolodner RD, Vogelstein B, Kunkel TA, Baylin SB. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A 1998;95:6870-6875.
86. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, Jessup JM, Kolodner R. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 1997;57:808-811.
87. Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD, Markowitz SD. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95:8698-8702.
88. Catteau A, Harris WH, Xu CF, Solomon E. Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. Oncogene 1999;18:1957-1965.
89. Rice JC, Ozcelik H, Maxeiner P, Andrulis I, Futscher BW. Methylation of the BRCA1 promoter is associated with decreased BRCA1 mRNA levels in clinical breast cancer specimens. Carcinogenesis 2000;21:1761-1765.
90. Ohtani-Fujita N, Dryja TP, Rapaport JM, Fujita T, Matsumura S, Ozasa K, Watanabe Y, Hayashi K, Maeda K, Kinoshita S, Matsumura T, Ohnishi Y, Hotta Y, Takahashi R, Kato MV, Ishizaki K, Sasaki MS, Horsthemke B, Minoda K, Sakai T. Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. Cancer Genet Cytogenet 1997;98:43-49.
91. Kim MS, Lee J, Sidransky D. DNA methylation markers in colorectal cancer. Cancer Metastasis Rev 2010;29:181-206.
92. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009;59:225-249.
93. Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007;18:581-592.
94. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA Cancer J Clin 2009;59:366-378.
95. Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. Cancer Epidemiol Biomarkers Prev 2009;18:1688-1694.
96. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010.
97. Labianca R, Beretta GD, Kildani B, Milesi L, Merlin F, Mosconi S, Pessi MA, Prochilo T, Quadri A, Gatta G, de BF, Wils J. Colon cancer. Crit Rev Oncol Hematol 2010;74:106-133.
98. O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst 2004;96:1420-1425.
99. Waldner MJ, Neurath MF. The molecular therapy of colorectal cancer. Mol Aspects Med 2010;31:171-178.

100. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-767.
101. Lindblom A. Different mechanisms in the tumorigenesis of proximal and distal colon cancers. Curr Opin Oncol 2001;13:63-69.
102. Gervaz P, Buhler L, Scheiwiller A, Morel P. A tale of two colons and two cancers. Distinct carcinogenesis and clinical outcome according to location proximal or distal to the splenic flexure. Swiss Surg 2003;9:3-7.
103. Gervaz P, Bucher P, Morel P. Two colons-two cancers: paradigm shift and clinical implications. J Surg Oncol 2004;88:261-266.
104. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology 2010;138:2059-2072.
105. Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B. Mutations of mitotic checkpoint genes in human cancers. Nature 1998;392:300-303.
106. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 2008;135:1079-1099.
107. Perez dC, I, de CG, Malumbres M. A census of mitotic cancer genes: new insights into tumor cell biology and cancer therapy. Carcinogenesis 2007;28:899-912.
108. Sieber OM, Heinimann K, Tomlinson IP. Genomic instability--the engine of tumorigenesis? Nat Rev Cancer 2003;3:701-708.
109. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Powell SM, Jen J, Hamilton SR, . Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812-816.
110. Akiyama Y, Sato H, Yamada T, Nagasaki H, Tsuchiya A, Abe R, Yuasa Y. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. Cancer Res 1997;57:3920-3923.
111. Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, . Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75-80.
112. Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, Lee TY, Bodmer D, Hoenselaar E, Hendriks-Cornelissen SJ, Tsui WY, Kong CK, Brunner HG, van Kessel AG, Yuen ST, van Krieken JH, Leung SY, Hoogerbrugge N. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nat Genet 2009;41:112-117.
113. Niessen RC, Hofstra RM, Westers H, Ligtenberg MJ, Kooi K, Jager PO, de Groote ML, Dijkhuizen T, Olderode-Berends MJ, Hollema H, Kleibeuker JH, Sijmons RH. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. Genes Chromosomes Cancer 2009;48:737-744.
114. Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Res 2002;62:3925-3928.
115. Gylling AH, Nieminen TT, Abdel-Rahman WM, Nuorva K, Juhola M, Joensuu EI, Jarvinen HJ, Mecklin JP, Aarnio M, Peltomaki PT. Differential cancer predisposition in Lynch syndrome: insights from molecular analysis of brain and urinary tract tumors. Carcinogenesis 2008;29:1351-1359.
116. Hitchins M, Williams R, Cheong K, Halani N, Lin VA, Packham D, Ku S, Buckle A, Hawkins N, Burn J, Gallinger S, Goldblatt J, Kirk J, Tomlinson I, Scott R, Spigelman A, Suter C, Martin D, Suthers G, Ward R. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;129:1392-1399.
117. Hitchins MP, Wong JJ, Suthers G, Suter CM, Martin DI, Hawkins NJ, Ward RL. Inheritance of a cancer-associated MLH1 germ-line epimutation. N Engl J Med 2007;356:697-705.

118. Morak M, Schackert HK, Rahner N, Betz B, Ebert M, Walldorf C, Royer-Pokora B, Schulmann K, von Knebel-Doeberitz M, Dietmaier W, Keller G, Kerker B, Leitner G, Holinski-Feder E. Further evidence for heritability of an epimutation in one of 12 cases with MLH1 promoter methylation in blood cells clinically displaying HNPCC. Eur J Hum Genet 2008;16:804-811.
119. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004;36:497-501.
120. Valle L, Carbonell P, Fernandez V, Dotor AM, Sanz M, Benitez J, Urioste M. MLH1 germline epimutations in selected patients with early-onset non-polyposis colorectal cancer. Clin Genet 2007;71:232-237.
121. Niv Y. Microsatellite instability and MLH1 promoter hypermethylation in colorectal cancer. World J Gastroenterol 2007;13:1767-1769.
122. Issa JP, Shen L, Toyota M. CIMP, at last. Gastroenterology 2005;129:1121-1124.
123. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007;50:113-130.
124. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A 2007;104:18654-18659.
125. Ahlquist T, Lind GE, Costa VL, Meling GI, Vatn M, Hoff GS, Rognum TO, Skotheim RI, Thiis-Evensen E, Lothe RA. Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers. Mol Cancer 2008;7:94.
126. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-793.
127. Gazin C, Wajapeyee N, Gobeil S, Virbasius CM, Green MR. An elaborate pathway required for Ras-mediated epigenetic silencing. Nature 2007;449:1073-1077.
128. Nagasaka T, Koi M, Kloor M, Gebert J, Vilkin A, Nishida N, Shin SK, Sasamoto H, Tanaka N, Matsubara N, Boland CR, Goel A. Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950-60, 1960.
129. O'Brien MJ, Yang S, Mack C, Xu H, Huang CS, Mulcahy E, Amorosino M, Farraye FA. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. Am J Surg Pathol 2006;30:1491-1501.
130. O'Brien MJ. Hyperplastic and serrated polyps of the colorectum. Gastroenterol Clin North Am 2007;36:947-68, viii.
131. Ordway JM, Williams K, Curran T. Transcription repression in oncogenic transformation: common targets of epigenetic repression in cells transformed by Fos, Ras or Dnmt1. Oncogene 2004;23:3737-3748.
132. Velho S, Moutinho C, Cirnes L, Albuquerque C, Hamelin R, Schmitt F, Carneiro F, Oliveira C, Seruca R. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis? BMC Cancer 2008;8:255.
133. Rosenberg DW, Yang S, Pleau DC, Greenspan EJ, Stevens RG, Rajan TV, Heinen CD, Levine J, Zhou Y, O'Brien MJ. Mutations in BRAF and KRAS differentially distinguish serrated versus non-serrated hyperplastic aberrant crypt foci in humans. Cancer Res 2007;67:3551-3554.
134. Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, Barker MA, Arnold S, McGivern A, Matsubara N, Tanaka N, Higuchi T, Young J, Jass JR, Leggett BA. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137-1144.

135. Jass JR, Whitehall VL, Young J, Leggett BA. Emerging concepts in colorectal neoplasia. Gastroenterology 2002;123:862-876.
136. Patra SK. Ras regulation of DNA-methylation and cancer. Exp Cell Res 2008;314:1193-1201.
137. Esteller M, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Watkins DN, Issa JP, Sidransky D, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res 2000;60:2368-2371.
138. Wolf P, Hu YC, Doffek K, Sidransky D, Ahrendt SA. O(6)-Methylguanine-DNA methyltransferase promoter hypermethylation shifts the p53 mutational spectrum in non-small cell lung cancer. Cancer Res 2001;61:8113-8117.
139. Menigatti M, Pedroni M, Verrone AM, Borghi F, Scarselli A, Benatti P, Losi L, Di GC, Schar P, Marra G, Ponz de LM, Roncucci L. O6-methylguanine-DNA methyltransferase promoter hypermethylation in colorectal carcinogenesis. Oncol Rep 2007;17:1421-1427.
140. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG, Buckmeier J, Alberts DS, Hamilton SR, Issa JP. MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 2005;97:1330-1338.
141. Wajapeyee N, Serra RW, Zhu X, Mahalingam M, Green MR. Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. Cell 2008;132:363-374.
142. Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, Whitehall VL, Leggett BA, Laird PW. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PLoS One 2009;4:e8357.
143. Mooi WJ, Peeper DS. Oncogene-induced cell senescence--halting on the road to cancer. N Engl J Med 2006;355:1037-1046.
144. Coebergh JW, Reedijk AM, de VE, Martos C, Jakab Z, Steliarova-Foucher E, Kamps WA. Leukaemia incidence and survival in children and adolescents in Europe during 1978-1997. Report from the Automated Childhood Cancer Information System project. Eur J Cancer 2006;42:2019-2036.
145. Kaatsch P. Epidemiology of childhood cancer. Cancer Treat Rev 2010;36:277-285.
146. Pieters R, Schrappe M, De LP, Hann I, De RG, Felice M, Hovi L, LeBlanc T, Szczepanski T, Ferster A, Janka G, Rubnitz J, Silverman L, Stary J, Campbell M, Li CK, Mann G, Suppiah R, Biondi A, Vora A, Valsecchi MG. A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet 2007;370:240-250.
147. Pieters R. Infant acute lymphoblastic leukemia: Lessons learned and future directions. Curr Hematol Malig Rep 2009;4:167-174.
148. Greaves MF. Infant leukaemia biology, aetiology and treatment. Leukemia 1996;10:372-377.
149. Jansen MW, Corral L, van der Velden VH, Panzer-Grumayer R, Schrappe M, Schrauder A, Marschalek R, Meyer C, den Boer ML, Hop WJ, Valsecchi MG, Basso G, Biondi A, Pieters R, van Dongen JJ. Immunobiological diversity in infant acute lymphoblastic leukemia is related to the occurrence and type of MLL gene rearrangement. Leukemia 2007;21:633-641.
150. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, Sallan SE, Lander ES, Golub TR, Korsmeyer SJ. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet 2002;30:41-47.

151. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002;1:133-143.
152. Stumpel DJ, Schneider P, van Roon EH, Boer JM, De LP, Valsecchi MG, de Menezes RX, Pieters R, Stam RW. Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options. Blood 2009;114:5490-5498.
153. Gu Y, Nakamura T, Alder H, Prasad R, Canaani O, Cimino G, Croce CM, Canaani E. The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell 1992;71:701-708.
154. Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. Nat Genet 1992;2:113-118.
155. Tkachuk DC, Kohler S, Cleary ML. Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell 1992;71:691-700.
156. Ziemin-van der Poel S, McCabe NR, Gill HJ, Espinosa R, III, Patel Y, Harden A, Rubinelli P, Smith SD, LeBeau MM, Rowley JD, . Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. Proc Natl Acad Sci U S A 1991;88:10735-10739.
157. Slany RK. The molecular biology of mixed lineage leukemia. Haematologica 2009;94:984-993.
158. Hanson RD, Hess JL, Yu BD, Ernst P, van LM, Berns A, van der Lugt NM, Shashikant CS, Ruddle FH, Seto M, Korsmeyer SJ. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. Proc Natl Acad Sci U S A 1999;96:14372-14377.
159. Ingham PW. trithorax and the regulation of homeotic gene expression in Drosophila: a historical perspective. Int J Dev Biol 1998;42:423-429.
160. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. Nature 1995;378:505-508.
161. Breen TR, Harte PJ. Molecular characterization of the trithorax gene, a positive regulator of homeotic gene expression in Drosophila. Mech Dev 1991;35:113-127.
162. Hess JL, Yu BD, Li B, Hanson R, Korsmeyer SJ. Defects in yolk sac hematopoiesis in Mll-null embryos. Blood 1997;90:1799-1806.
163. Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, Komori T. Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. Blood 1998;92:108-117.
164. Ernst P, Fisher JK, Avery W, Wade S, Foy D, Korsmeyer SJ. Definitive hematopoiesis requires the mixed-lineage leukemia gene. Dev Cell 2004;6:437-443.
165. Hsieh JJ, Cheng EH, Korsmeyer SJ. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell 2003;115:293-303.
166. Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 2003;23:186-194.
167. Takeda S, Chen DY, Westergard TD, Fisher JK, Rubens JA, Sasagawa S, Kan JT, Korsmeyer SJ, Cheng EH, Hsieh JJ. Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. Genes Dev 2006;20:2397-2409.
168. Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M. Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. Blood 2002;100:3710-3718.

169. Hess JL. MLL: a histone methyltransferase disrupted in leukemia. Trends Mol Med 2004;10:500-507.
170. Ernst P, Wang J, Korsmeyer SJ. The role of MLL in hematopoiesis and leukemia. Curr Opin Hematol 2002;9:282-287.
171. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 2007;7:823-833.
172. Popovic R, Zeleznik-Le NJ. MLL: how complex does it get? J Cell Biochem 2005;95:234-242.
173. Zeleznik-Le NJ, Harden AM, Rowley JD. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. Proc Natl Acad Sci U S A 1994;91:10610-10614.
174. Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F, Slany RK. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res 2002;30:958-965.
175. Bach C, Mueller D, Buhl S, Garcia-Cuellar MP, Slany RK. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. Oncogene 2009;28:815-823.
176. Yokoyama A, Somervaille TC, Smith KS, Rozenblatt-Rosen O, Meyerson M, Cleary ML. The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. Cell 2005;123:207-218.
177. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. Nature 1999;399:491-496.
178. Mujtaba S, Zeng L, Zhou MM. Structure and acetyl-lysine recognition of the bromodomain. Oncogene 2007;26:5521-5527.
179. Santillan DA, Theisler CM, Ryan AS, Popovic R, Stuart T, Zhou MM, Alkan S, Zeleznik-Le NJ. Bromodomain and histone acetyltransferase domain specificities control mixed lineage leukemia phenotype. Cancer Res 2006;66:10032-10039.
180. Ernst P, Wang J, Huang M, Goodman RH, Korsmeyer SJ. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol Cell Biol 2001;21:2249-2258.
181. Briggs SD, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SY, Winston F, Allis CD. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev 2001;15:3286-3295.
182. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 2002;10:1107-1117.
183. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell 2002;10:1119-1128.
184. Richardson C, Jasin M. Frequent chromosomal translocations induced by DNA double-strand breaks. Nature 2000;405:697-700.
185. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 2007;7:823-833.
186. Collins EC, Pannell R, Simpson EM, Forster A, Rabbitts TH. Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. EMBO Rep 2000;1:127-132.
187. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y. hDOT1L links histone methylation to leukemogenesis. Cell 2005;121:167-178.
188. Zhang W, Xia X, Reisenauer MR, Hemenway CS, Kone BC. Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCalpha in an aldosterone-sensitive manner. J Biol Chem 2006;281:18059-18068.

189. Bitoun E, Oliver PL, Davies KE. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum Mol Genet 2007;16:92-106.
190. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. Cancer Res 2005;65:11367-11374.
191. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia X, Jesneck J, Bracken AP, Silverman LB, Kutok JL, Kung AL, Armstrong SA. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14:355-368.
192. de WE, van SB. Chromatin domains in higher eukaryotes: insights from genome-wide mapping studies. Chromosoma 2009;118:25-36.

Chapter 2

Tumour-specific methylation of *PTPRG* intron 1 locus in sporadic and Lynch syndrome colorectal cancer

Eur J Hum Genet (2011) 19(3):307-12

Eddy H.J. van Roon[1-2,&], Noel F.C.C. de Miranda[1], Merlijn P. van Nieuwenhuizen[1], Emile J. de Meijer[2], Marjo van Puijenbroek[1], Pearlly S. Yan[3], Tim H.-M. Huang[3], Tom van Wezel[1], Hans Morreau[1*], Judith M. Boer[2*]

[1]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands ; [2]Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands ; [3]Molecular Virology, Immunology, and Medical Genetics, Division of Human Cancer Genetics, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA; [&]Current address: Department of Pediatric Oncology and Hematology, Erasmus Medical Center-Sophia Children's Hospital, Rotterdam, The Netherlands.

TUMOUR-SPECIFIC METHYLATION OF *PTPRG* INTRON 1 LOCUS IN SPORADIC AND LYNCH SYNDROME COLORECTAL CANCER

Abstract

DNA methylation is a hallmark in a subset of right-sided colorectal cancers. Methylation-based screening may improve prevention and survival rate for this type of cancer, which is often clinically asymptomatic in the early stages. We aimed to discover prognostic or diagnostic biomarkers for colon cancer by comparing DNA methylation profiles of right-sided colon tumours and paired normal colon mucosa using an 8.5k CpG island microarray. We identified a diagnostic CpG-rich region, located in the first intron of the *PTPRG* gene, with altered methylation already in the adenoma stage, i.e. prior to the carcinoma transition. Validation of this region in an additional cohort of 103 sporadic colorectal tumours and 58 paired normal mucosa tissue samples showed 94% sensitivity and 96% specificity. Interestingly, comparable results were obtained when screening a cohort of Lynch syndrome-associated cancers. Functional studies showed that *PTPRG* intron 1 methylation did not directly affect *PTPRG* expression, however, the methylated region overlapped with a binding site of the insulator protein CTCF. Chromatin immunoprecipitation (ChIP) showed that methylation of the locus was associated with absence of CTCF binding. Methylation-associated changes in CTCF binding to *PTPRG* intron 1 could have implications on tumour gene expression by enhancer blocking, chromosome loop formation or abrogation of its insulator function. The high sensitivity and specificity for the *PTPRG* intron 1 methylation in both sporadic and hereditary colon cancers support biomarker potential for early detection of colon cancer.

Introduction

DNA methylation is a common mechanism in colorectal tumorigenesis[1, 2]. Over the last decade, several genome-wide array-based methods have been developed, allowing the discovery of novel tumour-specific methylated loci. Enzymatic (HELP[3], MMASS[4], DMH[5], CHARM[6]) and ChIP methods[7] are most commonly used for genome-wide screening of DNA methylation, in combination with CpG island or promoter microarrays. An alternative genome-wide approach to identify genes silenced by DNA methylation detects expression differences in cell lines treated with DNA demethylating agents[8, 9]. More recently, captured methylated DNA[10] and bisulphite-converted reduced representations[11] are analyzed by high-throughput sequencing strategies. The unbiased approaches have indicated that transcription regulation associated with CpG methylation is not restricted to promoter CpG islands[12, 13]. Conserved regions up to 2kb distant from the promoter, annotated as CpG island shores[12], and promoter CpG islands of lesser density, annotated as intermediate-CpG islands[13], undergo cancer-specific methylation more often than traditional promoter CpG islands. The methylation status of these regions is strongly related to gene expression and might have been underestimated in previous studies.

The aim of this study was to discover novel tumour-specific DNA methylation markers

in right-sided colon cancer. These tumours have a higher frequency of the CpG island methylator phenotype (CIMP). Additionally, right-sided tumours are often clinically asymptomatic at early stages, thus, patients would greatly benefit from a reliable screening method. We employed differential methylation hybridization (DMH) combined with a 8.5k CpG clone library microarray for the initial identification of differential methylation in a cohort of colon cancers[14]. This library is enriched for CG-rich areas throughout the genome, encompassing promoter CpG islands as well as CpG-rich island shores and intermediate-CpG islands[15]. We report tumour-specific methylation of the first intron of the receptor protein-tyrosine phosphatase gamma gene (*PTPRG*), in both sporadic and Lynch syndrome colorectal cancers. Additionally we demonstrate that methylation of this region affects its binding to the CCCTC-binding factor (zinc finger protein, CTCF).

Materials and Methods

Tissue

Anonymized samples were obtained from patients who underwent surgery between 1988 and 2006 at the Leiden University Medical Center (Leiden, The Netherlands) or at the Rijnland Hospital (Leiderdorp, The Netherlands). Tumour sections were micro-dissected to minimize normal epithelium and stromal cells. DNA was isolated from fresh-frozen tissue using a previously described method[16], and from formalin-fixed paraffin-embedded (FFPE) tissue using the Wizard Genomic DNA Purification kit (Promega, Madison, WI, USA). We used available normal mucosa from the same individuals as control to correct for age-dependent methylation. Age, location and microsatellite instability (MSI) status for the sporadic tumours are listed in Supplementary Table S1, and for the Lynch syndrome-associated tumours in Supplementary Table S2. The colorectal cancer cell lines SW48, RKO, SW480, Caco2, SW837, and LS411 were obtained from the American Type Culture Collection (Manassas, VA, USA). DNA was isolated from these cell lines as described previously[16]. RNA was isolated using TRIZOL (Invitrogen, Carlsbad, CA, USA) and subsequently purified with Qiagen RNeasy columns combined with the RNase-free DNase kit (Qiagen Sciences, Germantown, MD, USA). The present study was approved by the Medical Ethics committee of the LUMC (protocol P01-019). Cases were analyzed following the medical ethical guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences.

CpG island microarrays

CpG island clone inserts (8544) were amplified using vector-based primers as described previously[14, 17]. The CpG island clone library was originally generated by the Sanger Centre from affinity-purified *in vitro* methylated DNA fragments[15]. Clone sequence information was downloaded from the Toronto Microarray Facility. PCR products were spotted onto CodeLink (GE Healthcare, Munich, Germany) slides using an OmniGrid arrayer (Genomic Solutions, Ann Arbor, MI, USA) at the Leiden Genome Technology Center (www.lgtc.nl) as described[18].

Differential methylation hybridization

DMH was performed according to Yan *et al.*[14]. Cy5-labeled amplicons, representing methylated DNA fragments derived from tumours and paired normal mucosa samples,

were co-hybridized to the CpG island microarrays with a Cy3-labeled common reference amplicon consisting of a pool of DNA from the six colorectal cancer cell lines described above. Detection was done on a G2565BA scanner (Agilent Technologies, Santa Clara, CA, USA) and image analysis using GenePix6.0 (Molecular Devices, Union City, CA, USA). Pre-processing, including normalization, was performed using the GenePix error model in Rosetta Resolver version 5.0 (Rosetta Biosoftware, Seattle, WA, USA). Microarray data for the tumour and normal samples were compared using an error-weighted ANOVA model and corrected for multiple testing[19] in Rosetta Resolver. Microarray data are available on Gene Expression Omnibus with accession number GSE21181.

Bisulphite sequence analysis (BSA)

DNA samples (500 ng) were converted using the EZ DNA methylation Gold bisulphite kit (Zymo Research). Primers (Supplementary Table S3) for BSA of ten CpGs in the *PTPRGint1* locus (Supplementary Figure S1) were designed using MethPrimer.[20] Amplification was carried out in a DNA Engine Dyad Peltier Thermal Cycler (Bio-Rad, Hercules, CA, USA) using AmpliTaq Gold PCR buffer and enzyme (Applied Biosystems, Foster City, USA). For the direct BSA, PCR products were sequenced using the right primer, resulting in nine interpretable CpGs (CpG2-10). For additional clonal BSA, *PTPRGint1* PCR fragments were cloned into TOP10 E. coli bacteria using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA, USA). Sequence alignment analysis was performed using ClustalW[21] and BioEdit.[22] CpG dinucleotides in the direct BSA were scored as being methylated when the ratio of the cytosine/thymine peaks was above 0.4. The BiQ analyzer software was used for visualization.[23]

MS-MLPA assay

Custom MS-MLPA probes (Supplementary Table S3) for the *PTPRGint1* locus were designed in Primer3[24] and included the HhaI site in CpG9 (see Figure 1B). As a control we used a *BRCA2* probe set from the SALSA MS-MLPA KIT ME001B Tumour suppressor-1 kit (MRC-Holland, Amsterdam, The Netherlands). Fragment analysis was performed on an ABI 3130 (Applied Biosystems, Foster City, US). The MS-MPLA was performed as described using 50 ng of genomic DNA[25]. A negative, unmethylated control (human semen DNA) and a 100% methylated DNA control (CpGenome Universal methylated DNA, Chemicon, Millipore, Billerica, MA, USA), were included in every experiment to asses HhaI cleavage and PCR. Fragment analysis was performed in GeneMapper (Applied Biosystems). *PTPRGint1* peak heights were normalized by comparison with the *BRCA2* peak height from the same run. Subsequently, the *PTPRGint1*/*BRCA2* ratio from the digested reaction was divided by the *PTPRGint1*/*BRCA2* ratio from the undigested reaction resulting in one ratio per sample that represented the fraction of methylated CpG9. Ten independent measurements provided a ratio distribution for unmethylated (mean 0.108, sd 0.037) and fully methylated (mean 0.833, sd 0.148) control DNA. Samples were typed as being unmethylated or methylated when they were within three standard deviations of the mean of the unmethylated and methylated reference samples, respectively. Samples with ratios in between these standard deviation boundaries were scored as partially methylated. For specificity and sensitivity calculations, partially methylated samples were considered methylated. Determination of specificity between normal and tumour tissue was performed by a Chi-square test.

Real-time RT-PCR

cDNA was generated using the random primer protocol from the Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Science, Indianapolis, IN, USA) using 1 µg of RNA. Intron-spanning primers (Supplementary Table S3) were designed to target exon 1 and 2 of the main *PTPRG* transcript in Primer3[24]. Reactions were performed in duplicate on a Light Cycler 480 (Roche Applied Science) using IQ SYBR Green SuperMix (Biorad). High resolution melting curve analysis was performed to check primer specificity. Reactions with more than one peak in the melting curve were discarded, as were samples where the standard deviation between technical replicates was above one Ct value. A standard curve was generated using five 1:10 dilutions of pooled cDNA from colon cancer cell lines (SW48, RKO, SW480, Caco2, SW837, and LS411), showing an efficiency of 100%. The Ct values that were used for analysis were between 23 and 33. Relative mRNA concentrations were calculated from this standard curve. Stably expressed control genes for normalization were selected with the GeNorm applet; the two most stably expressed genes were used for normalization of each tumour cDNA (*CPSF6* and *EEF1*)[26].

CTCF ChIP and quantitative PCR

The primary colon cancer cell line KP7038t, established at the Department of Pathology at the LUMC, was grown in GIBCO RPMI 1640 with glutamax (Invitrogen, Carlsbad, CA, USA), 10% foetal calf serum, and penicillin/streptomycin (50µg/ml). Tumour-associated fibroblasts (KP7038f), collected from the primary tumour, as well as cell lines RKO and SW480 were grown under identical conditions. Cells at approximately 80% confluency were fixed with 1% formaldehyde for 10 minutes at room temperature. The formaldehyde was quenched with glycine (0.125 M) and the cells were harvested by scraping. Chromatin immunoprecipitations were performed using the SimpleChIP Enzymatic Chromatin IP kit (Cell Signaling Technology, Danvers, MA, US) with 10 µl anti-CTCF antibody (D31H2 XP; Cell Signaling Technology). Normal colon mucosa was collected from a patient who underwent a colon colonoscopy unrelated to colon cancer. Twenty 50µm sections, cut from macro-dissected frozen tissue, were fixed in PBS with 1% formaldehyde. After quenching, the tissue was micro-dissected and processed for ChIP followed by duplicate PCR reactions as above. Primers (Supplementary Table S3) targeting control regions were selected from Kim *et al*.[27]. The Ct values that were used for analysis were between 25 and 40. Standard curves were generated using four consecutive 1:5 dilutions of input DNA for both cultures (non-immunoprecipitated, cross-linked DNA) to determine relative DNA concentrations. For comparison between pull-downs, relative DNA concentrations of the CTCF and IgG antibody pull-downs were normalized with the corresponding relative concentration from the histone H3 antibody pull-down.

Results

Locus PTPRGint1 is methylated in colorectal adenomas and carcinomas

Methylation profile comparison by ANOVA of 15 carcinomas, three adenomas and eight paired normal mucosa samples identified 20 differentially methylated loci for the three tissue groups (false discovery rate < 0.01%). For all but one of these loci methylation in the adenomas was comparable to the normal samples while carcinomas showed increased methylation (Figure 1A). The most significant CpG island clone was 47B02 that showed increased methylation in both adenomas and carcinomas compared to normal mucosa. Therefore, we performed validation experiments for the corresponding locus, which mapped to the first intron of the *PTPRG* gene (chr3: 61524993-61525363, UCSC assembly: March 2006, see Supplementary Figure S1), referred to as *PTPRGint1*.

Analysis at single CpG resolution using direct BSA showed that 17 out of 18 tumour samples were fully methylated in the *PTPRGint1* region, while one carcinoma was unmethylated for CpG dinucleotides 8-10 (Figure 1B, bottom panel). In contrast, normal colon samples were mostly unmethylated (Figure 1B, upper panel). CpGs 7-10 were most informative to distinguish between tumour and normal in this set of samples. These results were confirmed at the single chromosome level using clonal BSA (Supplementary Figure S2). In summary, the microarray-based finding of differential *PTPRGint1* methylation in right-sided tumours was confirmed, and extended to left-sided adenomas and carcinomas.

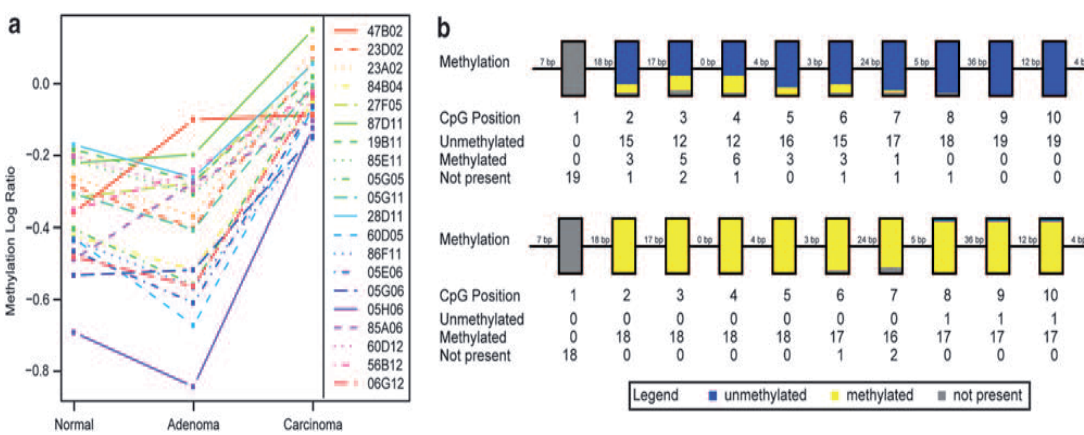


Figure 1- Identification and validation of the differentially methylated locus *PTPRGint1*. (A) Trend plot of the top-20 differentially methylated microarray clones (FDR ≤ 0.01%) showing the average log10 ratios in the normal, adenoma and carcinoma samples compared to the colorectal cancer cell line reference panel. Clone 47B02, corresponding to *PTPRGint1*, had a similarly increased log ratio in adenomas and carcinomas compared to normals (red solid line). (B) BiQ summary of direct BSA of *PTPRGint1* (CpG dinucleotides 1-10) in paired normal colon mucosa (n=19, upper panel) and colon tumours (n = 18, 12 right- and left-sided carcinomas and 6 adenomas, lower panel) showed highest specificity and sensitivity in the four most 3' CpGs measured. Each box corresponds to one CpG position in the genomic sequence, while colours summarize the methylation states of all sequenced samples at that position.

Methylation of PTPRGint1 CpG9 has high sensitivity and specificity

To assess the sensitivity and specificity of *PTPRGint1* methylation to discriminate between cancer and normal tissue, we developed a high-throughput MS-MLPA assay (Figure 2A, 2B). We tested an FFPE cohort of 103 tumours and 58 corresponding normal tissues, which allowed us to assess the possibility of age-related methylation often seen in aging mucosa. Of the 67 carcinoma samples 94% showed methylation of the targeted CpG dinucleotide (61 fully, two partially methylated), while 46 (95.8%) of the 48 corresponding normal samples were unmethylated, and the remaining two partially methylated (Figure 2C, Table 1). Comparing the methylation numbers between normal and tumour tissue by Chi-square test provided a highly significant p-value of $9.8x10^{-110}$. *PTPRGint1* methylation was independent of MSI status in the sporadic carcinomas, as both MSI-High and stable cases were methylated. A relatively small group of 18 sporadic adenomas (13 low grade dysplastic and five high grade dysplastic) and neighbouring normal mucosa of 10 of these showed *PTPRGint1* methylation in all adenomas, but not in the available normal mucosa (Figure 2C, Table 1). To assess whether *PTPRGint1* methylation is an early event in colorectal carcinogenesis, several colon lesions preceding the adenoma/carcinoma stages were studied. All six hyperplastic polyps and 11 out of 12 serrated adenomas tested showed *PTPRGint1* methylation (Figure 2C).

Table 1: Sensitivity and specificity of the *PTPRGint1* locus CpG9 methylation in sporadic tumours and tumours associated with a specific MMR mutation

CpG9	Sensitivity Methylated tumours / Total Tumours	Specificity [1] Unmethylated normals / Total normals
Sporadic adenomas	100% (18/18)	100% (10/10)
Sporadic carcinomas	94% (63/67)	95.8% (46/48)
MLH1 mutated	100% (14/14)	87% (20/23)
MSH2 mutated	96% (18/19)	100% (24/24)
MSH6 mutated	86.7% (26/30)	100% (14/14)
Total Lynch	92.1% (58/63)	95.4% (62/65)

[1]The mutational analysis of the MMR genes was incomplete for four unpaired normal samples in the Lynch syndrome cohort. Therefore, these were only included in the total specificity calculations.

PTPRGint1 CpG9 methylation in Lynch syndrome associated colorectal cancer

The initial cohort for MS-MLPA validation contained two Lynch syndrome associated colorectal carcinomas that were both methylated. Therefore, we further investigated *PTPRGint1* methylation in 63 carcinomas from patients with a DNA mismatch repair (MMR) gene mutation (14 *MLH1* mutations, 19 *MSH2* mutations and 30 *MSH6* mutations). Of these 92.1% showed methylation (58 fully, 3 partially methylated), while 62 of the 65 (95.7%) corresponding normal samples were unmethylated (Figure 2C, Table 1; Chi-square p-value $3.3x10^{-153}$). In conclusion, methylation of *PTPRGint1* CpG9 has similarly high sensitivity and specificity in Lynch syndrome associated colorectal carcinomas as in sporadic colorectal cancer.

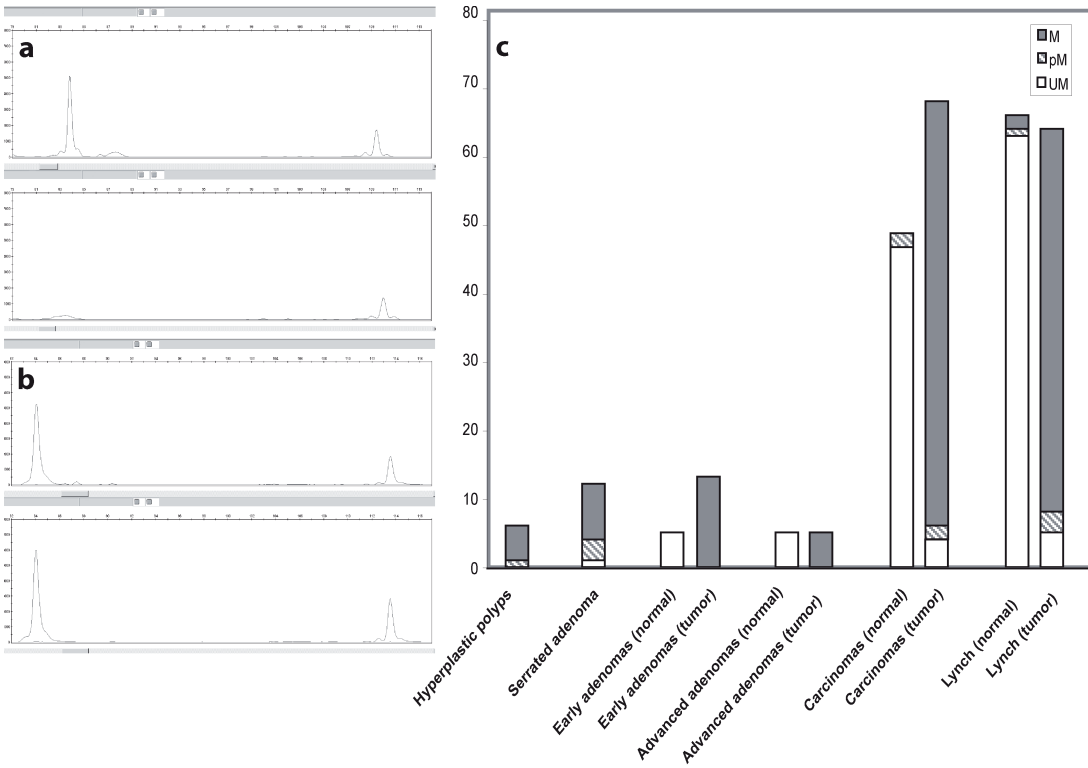


Figure 2 - *PTPRGint1* methylation detected by MS-MLPA. (a) GeneMapper output of the custom MS-MLPA to analyze *PTPRGint1* CpG9 methylation in a normal mucosa sample. Upper panel: The *PTPRGint1* peak was located at ~84 bp, the control peak was located at ~110 bp. Lower panel: Loss of *PTPRGint1* signal after HhaI digestion indicated an unmethylated CpG9. (b) GeneMapper output of the custom *PTPRGint1* MS-MLPA in the corresponding colon cancer sample. Upper panel: Undigested. Lower panel: HhaI digestion. Retention of the *PTPRGint1* marker signal indicated protection against HhaI digestion by CpG9 methylation. (c) Frequency of *PTPRGint1* CpG9 methylation in precursor lesions (hyperplastic polyps and serrated adenoma), early- and advanced adenomas, carcinomas, and corresponding normal mucosal tissue for the latter three groups. The number of samples typed as methylated (dark), partially methylated (striped) and unmethylated (white) in the MS-MLPA assay is indicated on the y-axis.

No silencing of PTPRG gene expression in methylated samples

To assess if hypermethylation of *PTPRGint1* affected expression of the gene, relative mRNA levels of *PTPRG* were studied in 15 right-sided colon carcinomas, three adenomas and 18 corresponding normal mucosa samples. The *PTPRG* gene encodes four protein coding isoforms (Source: HGNC Symbol; Acc:9671, aligned to Ensembl GRCh37). Using intron 1 spanning primers the full length transcripts ENST00000474889 and ENST00000295874 (missing one cassette exon) were analyzed (Figure 3). Two additional transcripts, ENST00000383711 and ENST00000394462 both starting at exon 3, were analyzed using intron 26 spanning primers and gave comparable results (data not shown). Expression of the *PTPRG* gene was detected in all samples. To assess the effects of *PTPRGint1* methylation on *PTPRG* expression, we compared the MS-MLPA methylation ratio with the *PTPRG* relative expression (Figure 3). We found that *PTPRG* was expressed at similar levels in the tumour and normal samples independent of methylation status. One patient (No. 28) showed increased *PTPRG* expression in both normal and tumour samples, thought to reflect individual expression differences and a possible copy number effect in the tumour.

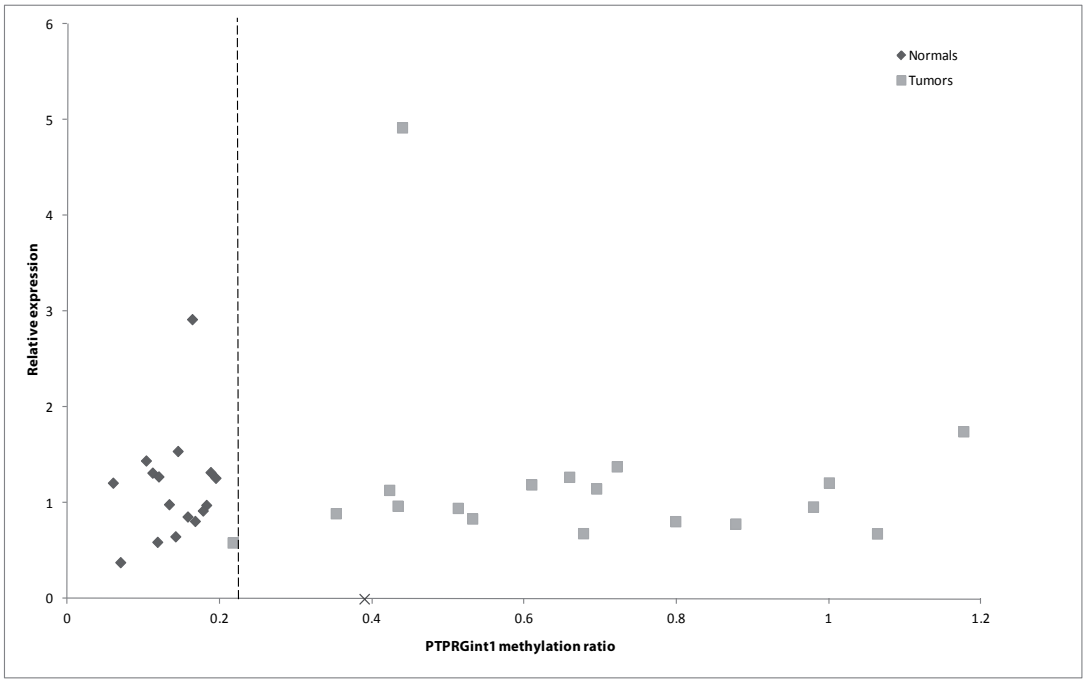
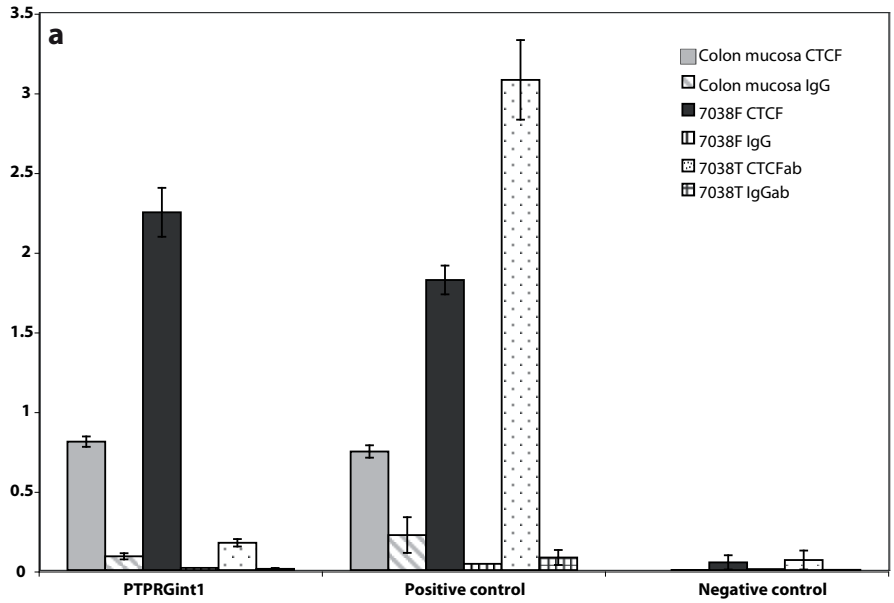


Figure 3 - Scatter plot of the relative *PTPRG* expression against the *PTPRGint1* methylation ratio according to the MS-MLPA. The dotted vertical line at 0.22 indicates the cut-off for unmethylated samples. Plotted data is representative of two independent experiments.

PTPRGint1 is a methylation-dependent CTCF binding site

The *PTPRGint1* locus overlapped with an experimentally defined CTCF binding site from a CTCF ChIP-chip study in fibroblasts[27] displayed in the genome web browsers USCS[28] and Ensembl[29]. This CTCF binding region of 750bp (OREG0015647; chr3: 61525101-61525851, UCSC assembly: March 2006) has a 262 bp overlap with the 3' part of the *PTPRGint1* locus (Supplementary Figure S1). We studied CTCF binding by ChIP in the primary tumour culture KP7038t that carried *PTPRGint1* methylation (MS-MLPA ratio 0.97, data not shown), unmethylated tumour-derived fibroblasts KP7038f from the same patient (MS-MLPA ratio 0.07, data not shown), as well as normal colon mucosa. We detected CTCF binding to *PTPRGint1* in the normal mucosa and KP7038f fibroblasts, but little binding to the primary tumour cells (Figure 4A).

To control for possible differences in the amount of CTCF protein between the samples, we compared the qPCR results for the *PTPRGint1* region with a positive control locus 7.9 Mb distant from *PTPRGint1* that was shown to bind CTCF and does not contain a CpG in its putative 20 bp consensus sequence[27]. This locus was enriched in pull-downs of all samples, including the primary tumour cells, indicating that lack of CTCF binding to *PTPRGint1* was not due to lack of CTCF protein. CTCF binding to *PTPRGint1* was comparable to the positive control in both KP7038f fibroblasts (ratio 1.2) and normal colon mucosa (ratio 1.1). However, the *PTPRGint1*/positive control ratio was 0.06 in the primary tumour culture KP7038t (Figure 4B). CTCF binding to *PTPRGint1* was similarly decreased in colorectal cancer cell lines RKO and SW480 (Figure 4B). These results indicate a significant decrease of bound CTCF to the methylated *PTPRGint1* region in tumour cells.



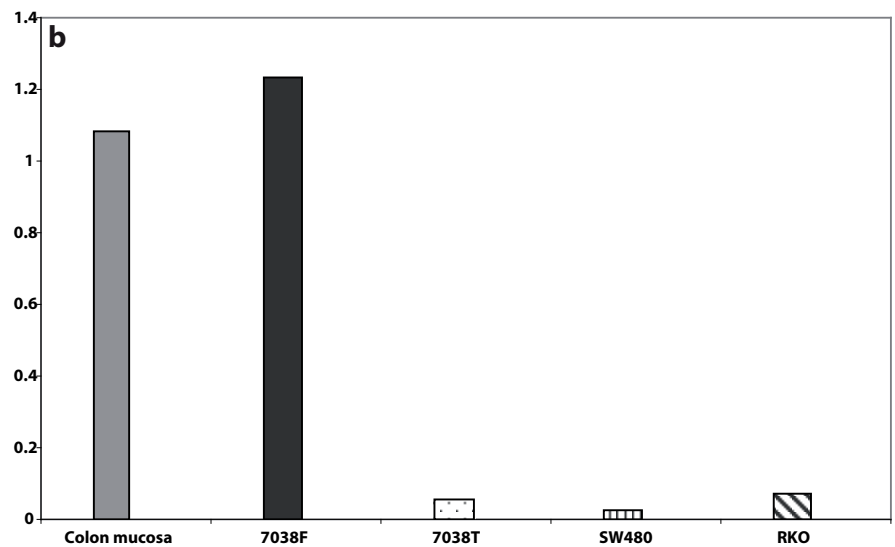


Figure 4 - (a) CTCF binding to *PTPRGint1*, positive- and negative control regions in normal colon mucosa (light grey), KP7038f (black) and KP7038t (dotted). The histone H3 normalized values of the CTCF antibody and IgG negative control antibody pull downs are given per primer pair. Standard errors represent the variability of duplicate PCR reactions. This is a representative experiment of three ChIPs. (b) The *PTPRGint1*/positive control ratio for normal colon mucosa, KP7038f, KP7038t (also shown in A), SW480 (vertically striped) and RKO (diagonally striped).

Discussion

We describe colorectal tumour-specific methylation of a locus in the first intron of the putative tumour suppressor gene *PTPRG* in both proximal and distal carcinomas and adenomas, including Lynch syndrome tumours. For these high-risk individuals, who are advised to undergo regular colonoscopies, no molecular markers have been described so far[30]. Assuming that a successful faecal or blood DNA test for *PTPRGint1* methylation could be developed, this is a promising discovery that would aid the early detection of colorectal tumours, independent of their aetiology.

PTPRGint1 is not located in a promoter CpG island but in the first intron, about 3kb from the transcriptional start site. We did not find a relation between the methylation status of *PTPRGint1* and *PTPRG* expression, indicating that *PTPRGint1* methylation does not lead to loss of function of PTPRG as has been observed for mutations in colon cancer, and deletions in lung carcinomas and renal carcinoma cell lines[31,32]. However, the identification of a methylation-sensitive CTCF binding site overlapping with *PTPRGint1* suggests that tumour-specific methylation may have a more distant effect. Differential binding of the insulator protein CTCF could have a major influence on expression of distant genes through alternative loop formation, as has been observed in β-globin expression in mouse models[33]. A recent study has shown that loss of CTCF binding to a boundary region upstream of *CDKN2A* resulted in spreading of repressed chromatin and DNA methylation into the p16 promoter with sequential down-regulation of *p16* expression[34]. The same study described that loss of upstream CTCF binding resulted in promoter DNA methylation of *RASSF1* and *CDH1*[34]. Contradictory to the finding that CTCF binding abrogation was shown to be causative of heterochromatin spreading and DNA methylation[34,35], is the observation that DNA methylation of CTCF binding sites is suggested to regulate CTCF binding[36, 37]. Interestingly, aberrant DNA methylation that excludes CTCF binding to intronic regulatory DNA was shown to promote expression of an oncogene, *BCL6*, in B cell lymphomas[38]. Although the sequence of events is unknown, the age-related aberrant hypermethylation often seen in colon cancer hints towards the latter.

We excluded differential peptidic abundance of CTCF between tumour and normal samples by successfully amplifying a positive control CTCF binding site on all samples. It remains to be demonstrated if CTCF protein modifications and its cellular and nuclear distribution are maintained in all tumour cells, both of which can influence CTCF activity and the binding to specific regions. More insight into the role of aberrant DNA methylation of *PTPRGint1* in the aetiology of cancer requires a better understanding of whether aberrant CTCF binding is caused by inhibition of protein activity or by initial aberrant methylation of the CTCF binding site.

Hypermethylation of the CpG island in the *PTPRG* gene promoter has been previously described in cutaneous T-cell lymphomas, melanoma cell lines, and gastric cancer[39-41]. Transcriptional down-regulation was shown to be associated with *PTPRG* promoter methylation in the cutaneous T-cell lymphomas study[39]. This study used a similar microarray for identification of differential methylation as the present study. We did not find differential methylation of the *PTPRG* promoter region between normal and colon tumour samples on the microarray. Moreover, BSA of colorectal cancer cell lines showed that the *PTPRG* promoter region was unmethylated (R. van Doorn, personal communication). Therefore, we have no indication for upstream spreading of DNA methylation from the CTCF binding

region at *PTPRGint1* towards the promoter.

In conclusion, this study provided evidence for tumour-specific hypermethylation of a CTCF binding site located in the first intron of *PTPRG*. The high specificity and sensitivity imply a possible utility for *PTPRGint1* methylation in new or existing colon-specific methylation marker panels. Especially the high level of *PTPRGint1* methylation in Lynch syndrome associated colorectal tumours is unique and could prove to be a valuable addition. Methylation-dependent absence of CTCF binding to the *PTPRGint1* locus suggests a possible effect on chromatin density or conformation that could play a role in colon tumorigenesis[42].

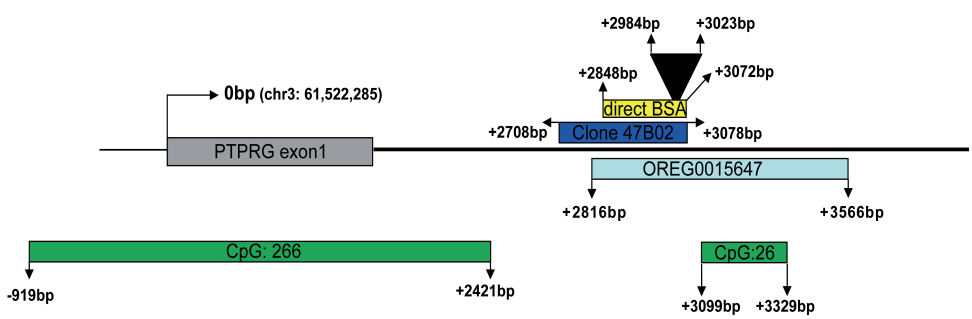
Acknowledgments

The authors thank Remco van Doorn and Wim Zoutman for sharing protocols and unpublished data and for technical assistance with the library amplification. The authors would also like to thank Jaap van Eendenburg for establishing primary culture KP7038. This study was supported by grants from the Stichting prof. A.A.H. Kassenaar fund and the Centre for Medical Systems Biology within the framework of the Netherlands Genomics Initiative (NGI)/ Netherlands Organization for Scientific Research (NWO).

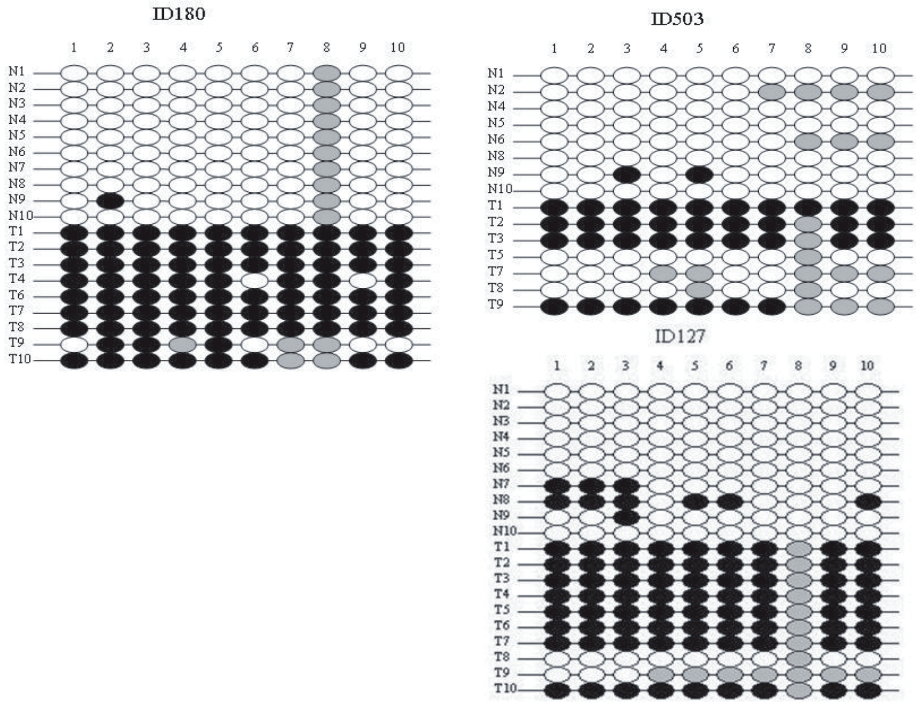
Conflict of interest

The authors (EHJvR, HM, and JB) have filed a patent application pertaining to intellectual property related in part to this report (P78692EP00 Submitted December 31, 2007). All other authors have nothing to disclose.

Supplementary data



Supplementary Figure S1 - Visual representation of the PTPRGint1 region. The *PTPRG* exon 1 (grey) and beginning of intron 1 (thick black line) are shown, with the transcription start site indicated with an arrow. Relative to the *PTPRG* start site, the following genomic features are shown: Promoter CpG island and intron 1 CpG island (green), ORegAnno CTCF fragment OREG0015647 (adapted from UCSC genome browser, light blue), CpG island microarray clone 47B02 (dark blue), amplified region by direct BSA (yellow). The black arrowhead indicates the approximate location of MS-MLPA probe hybridization. Visualization based on UCSC assembly: March 2006.



Supplementary Figure S2 - Clonal BSA of PTPRGint1 in three paired tumour-normal tissues**.** Visual representation of the clonal BSA results of ten CpG dinucleotides of PTPRGint1 in carcinomas ID503 and ID127 and adenoma ID180. Black dot: methylated CpG. White dot: unmethylated CpG. Grey dot: sequence not readable. In a partially methylated normal sample (ID127) not more than three out of ten alleles were found methylated for several CpGs. In addition, complete methylation of 9/9, 8/10 and 4/7 sequenced tumour alleles was found respectively, roughly corresponding to the estimated percentage of tumour cells observed in HE-stained tissue sections.

Supplementary Table S1 - Overview of the sporadic colorectal tumour samples and paired controls used in this study. Location, histology, age, sex and microsatellite instability status are listed, as well as the different techniques used on each sample. A gray box behind a sample in one of the last five columns indicates that the sample was tested with the technique listed in the header. Available at: http://www.nature.com/ejhg/journal/v19/n3/suppinfo/ejhg2010187s1.html?url=/ejhg/journal/v19/n3/full/ejhg2010187a.html

Supplementary Table S2 - Overview of Lynch cohort used. Location, histology, mismatch repair gene mutation, age and sex are listed for each sample as well as the different techniques used in this study. A gray box in the last column indicates that the sample was tested with the technique listed in the header. Available at: http://www.nature.com/ejhg/journal/v19/n3/suppinfo/ejhg2010187s1.html?url=/ejhg/journal/v19/n3/full/ejhg2010187a.html

Supplementary Table S3 Overview of the primers used in this study. Available at: http://www.nature.com/ejhg/journal/v19/n3/suppinfo/ejhg2010187s1.html?url=/ejhg/journal/v19/n3/full/ejhg2010187a.html

References

1. Grady WM, Carethers JM: Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 2008; **135**: 1079-1099.
2. Hiltunen MO, Alhonen L, Koistinaho J *et al*: Hypermethylation of the APC (adenomatous polyposis coli) gene promoter region in human colorectal carcinoma. *Int J Cancer* 1997; **70**: 644-648.
3. Khulan B, Thompson RF, Ye K *et al*: Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res* 2006; **16**: 1046-1055.
4. Ibrahim AE, Thorne NP, Baird K *et al*: MMASS: an optimized array-based method for assessing CpG island methylation. *Nucleic Acids Res* 2006; **34**: e136.
5. Huang TH, Perry MR, Laux DE: Methylation profiling of CpG islands in human breast cancer cells. *Hum Mol Genet* 1999; **8**: 459-470.
6. Irizarry RA, Ladd-Acosta C, Carvalho B *et al*: Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res* 2008; **18**: 780-790.
7. Weber M, Davies JJ, Wittig D *et al*: Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 2005; **37**: 853-862.
8. Schuebel KE, Chen W, Cope L *et al*: Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. *PLoS Genet* 2007; **3**: 1709-1723.
9. Suzuki H, Gabrielson E, Chen W *et al*: A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet* 2002; **31**: 141-149.
10. Down TA, Rakyan VK, Turner DJ *et al*: A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol* 2008; **26**: 779-785.
11. Meissner A, Mikkelsen TS, Gu H *et al*: Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 2008; **454**: 766-770.
12. Irizarry RA, Ladd-Acosta C, Wen B *et al*: The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009; **41**: 178-186.
13. Weber M, Hellmann I, Stadler MB *et al*: Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 2007; **39**: 457-466.
14. Yan PS, Chen CM, Shi H *et al*: Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. *Cancer Res* 2001; **61**: 8375-8380.
15. Cross SH, Charlton JA, Nan X *et al*: Purification of CpG islands using a methylated DNA binding column. *Nat Genet* 1994; **6**: 236-244.
16. Isola J, DeVries S, Chu L *et al*: Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples. *Am J Pathol* 1994; **145**: 1301-1308.
17. Yan PS, Efferth T, Chen HL *et al*: Use of CpG island microarrays to identify colorectal tumors with a high degree of concurrent methylation. *Methods* 2002; **27**: 162-169.
18. Knijnenburg J, Szuhai K, Giltay J *et al*: Insights from genomic microarrays into structural chromosome rearrangements. *Am J Med Genet A* 2005; **132**: 36-40.
19. Benjamini Y, Hochberg Y: Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 1995; **57**: 289-300.
20. Li LC, Dahiya R: MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 2002; **18**: 1427-1431.
21. Larkin MA, Blackshields G, Brown NP *et al*: Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; **23**: 2947-2948.

22. Tom Hall. BioEdit (URL: http://www.mbio.ncsu.edu/BioEdit/). 2008.
23. Bock C, Reither S, Mikeska T *et al*: BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* 2005; **21**: 4067-4068.
24. Rozen S, Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000; **132**: 365-386.
25. Nygren AO, Ameziane N, Duarte HM *et al*: Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Res* 2005; **33**: e128.
26. Vandesompele J, De PK, Pattyn F *et al*: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; **3**: RESEARCH0034.
27. Kim TH, Abdullaev ZK, Smith AD *et al*: Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 2007; **128**: 1231-1245.
28. Kuhn RM, Karolchik D, Zweig AS *et al*: The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res* 2009; **37**: D755-D761.
29. Hubbard TJ, Aken BL, Ayling S *et al*: Ensembl 2009. *Nucleic Acids Res* 2009; **37**: D690-D697.
30. Alemayehu A, Sebova K, Fridrichova I: Redundant DNA methylation in colorectal cancers of Lynch-syndrome patients. *Genes Chromosomes Cancer* 2008; **47**: 906-914.
31. LaForgia S, Morse B, Levy J *et al*: Receptor protein-tyrosine phosphatase gamma is a candidate tumor suppressor gene at human chromosome region 3p21. *Proc Natl Acad Sci U S A* 1991; **88**: 5036-5040.
32. Wang Z, Shen D, Parsons DW *et al*: Mutational analysis of the tyrosine phosphatome in colorectal cancers. *Science* 2004; **304**: 1164-1166.
33. Hou C, Zhao H, Tanimoto K *et al*: CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A* 2008; **105**: 20398-20403.
34. Witcher M, Emerson BM: Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell* 2009; **34**: 271-284.
35. Tiwari VK, Baylin SB: Breaching the boundaries that safeguard against repression. *Mol Cell* 2009; **34**: 395-397.
36. De La Rosa-Velazquez IA, Rincon-Arano H, itez-Bribiesca L *et al*: Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res* 2007; **67**: 2577-2585.
37. Xu J, Huo D, Chen Y *et al*: CpG island methylation affects accessibility of the proximal BRCA1 promoter to transcription factors. *Breast Cancer Res Treat* 2009.
38. Lai AY, Fatemi M, Dhasarathy A *et al*: DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *J Exp Med* 2010; **207**: 1939-1950.
39. van Doorn R, Zoutman WH, Dijkman R *et al*: Epigenetic profiling of cutaneous T-cell lymphoma: promoter hypermethylation of multiple tumor suppressor genes including BCL7a, PTPRG, and p73. *J Clin Oncol* 2005; **23**: 3886-3896.
40. Furuta J, Nobeyama Y, Umebayashi Y *et al*: Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. *Cancer Res* 2006; **66**: 6080-6086.
41. Wang JF, Dai DQ: Metastatic suppressor genes inactivated by aberrant methylation in gastric cancer. *World J Gastroenterol* 2007; **13**: 5692-5698.
42. Lewis A, Murrell A: Genomic imprinting: CTCF protects the boundaries. *Curr Biol* 2004; **14**: R284-R286.

Chapter 3

Early onset MSI-H colon cancer with *MLH1* promoter methylation, is there a genetic predisposition?

BMC Cancer (2010) 10:180

Eddy H.J. van Roon[1-2], Marjo van Puijenbroek[1], Anneke Middeldorp[1], Ronald van Eijk[1], Emile J. de Meijer[2], Dianhdra Erasmus[1], Kim A.D. Wouters[3], Manon van Engeland[3], Jan Oosting[1], Frederik J. Hes[2], Carli M.J. Tops[2], Tom van Wezel[1], Judith M. Boer[2-§], Hans Morreau[1-§]

[1]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.
[2]Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands.
[3]Department of Pathology, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands
[§]Corresponding authors

EARLY ONSET MSI-H COLON CANCER WITH *MLH1* PROMOTER METHYLATION, IS THERE A GENETIC PREDISPOSITION?

Abstract

Background: To investigate the etiology of *MLH1* promoter methylation in mismatch repair (MMR) mutation-negative early onset MSI-H colon cancer. As this type of colon cancer is associated with high ages, young patients bearing this type of malignancy are rare and could provide additional insight into the etiology of sporadic MSI-H colon cancer.
Methods: We studied a set of 46 MSI-H colon tumors cases with *MLH1* promoter methylation which was enriched for patients with an age of onset below 50 years (n=13). Tumors were tested for CIMP marker methylation and mutations linked to methylation: *BRAF, KRAS*, *GADD45A* and the *MLH1* -93G>A polymorphism. When available, normal colon and leukocyte DNA was tested for *GADD45A* mutations and germline *MLH1* methylation. SNP array analysis was performed on a subset of tumors.
Results: We identified two cases (33 and 60 years) with *MLH1* germline promoter methylation. *BRAF* mutations were less frequent in colon cancer patients below 50 years relative to patients above 50 years (p-value: 0.044). CIMP-high was infrequent and related to *BRAF* mutations in patients below 50 years. In comparison with published controls the G>A polymorphism was associated with our cohort. Although similar distribution of the pathogenic A allele was observed in the patients with an age of onset above and below 50 years, the significance for the association was lost for the group under 50 years. *GADD45A* sequencing yielded an unclassified variant. Tumors from both age groups showed infrequent copy number changes and loss-of-heterozygosity.
Conclusion: Somatic or germline *GADD45A* mutations did not explain sporadic MSI-H colon cancer. Although germline *MLH1* methylation was found in two individuals, locus-specific somatic *MLH1* hypermethylation explained the majority of sporadic early onset MSI-H colon cancer cases. Our data do not suggest an intrinsic tendency for CpG island hypermethylation in these early onset MSI-H tumors other than through somatic mutation of *BRAF*.

Background

High frequency of microsatellite instability (MSI-H) is the hallmark of tumors with a mismatch DNA repair (MMR) deficiency. This deficiency leads to an accumulation of somatic mutations, especially in repetitive coding or non-coding DNA sequences (microsatellites) in the genome. MSI-H in colon cancer is found in the context of Lynch syndrome, previously known as hereditary non-polyposis colorectal cancer (HNPCC), in which germline mutations in one of four mismatch repair genes (primarily in *MLH1* and *MSH2* [1] and to a lesser extent in *MSH6*[2], *PMS2*[3] or deletions in *EPCAM/TACSTD1* (leading to MSH2 methylation)[4, 5] are found. Approximately 15% of cases are due to somatic biallelic or hemiallelic DNA methylation of the CpG-rich *MLH1* promoter sequence, which is associated with gene silencing[6]. Colon cancers with sporadic MSI-H are observed more frequently in females and are often located proximal to the splenic flexure[7].
A clear association between increased age and occurrence of sporadic MSI-H colon

cancer was described in 2002 by Young et al.[8]. The combination of age at diagnosis and three pathological features (tumor heterogeneity, peritumoral lymphocytes and tumor-infiltrating lymphocytes) allowed positive identification of 94.5% of MSI-H cancers as either Lynch syndrome or sporadic[8]. As normal aging colon mucosa shows global hypomethylation and specific hypermethylation of tumor associated genes, this epigenetic accumulation can explain the association between sporadic MSI-H colon cancer and older age[9-11]. The rare sporadic cases diagnosed at a relatively young age can provide insight into the etiology of sporadic MSI-H. As young patients with sporadic MSI-H colon cancer are subjected to *MLH1* methylation without this age-associated epigenetic accumulation, a defect of DNA methylation maintenance or direct targeting of *MLH1* for methylation could be expected.

MLH1 methylation has been one of the hallmarks of the CpG island methylator phenotype (CIMP) since the phenotype was first described in 1999[12]. The high levels of methylation found in CIMP–high colon tumors suggest that a causative genetic or epigenetic defect influences the spread and initiation of methylation[13]. Somatic *BRAF* mutations, *MLH1* methylation and sporadic MSI-H are associated with CIMP-positive (CIMP-high and CIMP-low combined) colon tumors in which the bulk of aberrant methylation can be found[13, 14]. Aberrant methylation in CIMP-high tumors is thought to arise through an increase in *de novo* methylation. *KRAS* mutations have also been associated with elevated levels of aberrant DNA methylation, although discrepancies between marker panels and techniques showed variable levels of methylation. In general, *KRAS* mutations are associated with CIMP-low (also annotated CIMP2) colon tumors, in which increased levels of aberrant methylation can be detected to some extent, but at lower levels than in the CIMP-high tumors[13, 14].

The underlying causes leading to *MLH1* promoter hypermethylation and subsequently to sporadic MSI-H colon cancer are still largely unknown. A relatively new concept in the field of genetics is germline epimutation. Although rare, multiple studies have described inherited and *de novo* germline methylation of *MLH1* in patients with Lynch-like colon cancer[15-21]. Cases with confirmed or probable *MLH1* epimutations are documented to have the same range of tumors as described in Lynch syndrome patients, predominantly early-onset MSI colorectal cancer and endometrial cancer. Although possible, inheritance of the *MLH1* epimutation is described as very weak, as the *MLH1* epimutation is unstable in the germline[17, 18, 20, 21]. Paradoxically, patients suspected of having a genetic disorder based on a strong family history may be less likely to carry an epimutation[18]. Germline epimutations are thus highly suspected in young patients presenting with an MSI tumor without a clear family history. Increased risk of MSI-H tumors[22] and tumor-specific *MLH1* methylation[23] might also be associated with a single-nucleotide polymorphism (SNP) -93 bp from the *MLH1* transcription start site (rs1800734). This *MLH1* G>A polymorphism is associated with increased age of onset and CIMP and *BRAF* mutations in individuals with MSI-H tumors[24].

Another possible factor contributing to aberrant DNA methylation is inactivation of *GADD45A*[25], although this finding was later disputed[26, 27]. *GADD45A* (growth arrest and DNA-damage inducible protein 45 alpha) is a nuclear protein involved in maintenance of genomic stability, DNA repair and cell growth suppression[28, 29]. A recent publication has found *GADD45A* to be a key regulator of active DNA demethylation in *Xenopus* oocytes and cell lines through a DNA repair-induced mechanism[25]. Specific short interfering RNA (siRNA)-mediated knockdown of *GADD45A* and *GADD45B* in the colon cancer cell line RKO

induced hypermethylation of *MLH1, THBS1* and *p16*, three genes known to be involved in carcinogenesis of different types of tumors by DNA methylation[25].

In contrast to MSI-H colon cancer, chromosomal instability (CIN) is enhanced and more pronounced in tumors with a low frequency of microsatellite instability (MSS or MSI-L tumors). This relationship can also be deduced from the observation that MSS/MSI-L tumors often are aneuploid, whereas MSI-H tumors mostly are peri-diploid. Lynch syndrome-associated MSI-H colon cancer hardly shows chromosomal copy number alterations, and the few alterations are mainly restricted to copy neutral LOH (cnLOH) at the mutated locus, especially in *MLH1* mutated cases[30]. However, sporadic MSI-H colon cancer and MSI-H from patients with unclassified variants in MMR genes seem to show an enhanced (although subtle) number of chromosomal aberrations[30-33].

We studied 46 MSI-H colon tumors showing loss of *MLH1* expression and its heterodimer *PMS2* and methylation of the *MLH1* promoter. Pathogenic germline MMR mutation were excluded. We have primarily focused on comparing relatively young patients with patients of older ages to identify a possible cause for *MLH1* methylation in young individuals with colon cancer. Tumors were characterized for somatic *BRAF*, *KRAS*, *GADD45A* and the *MLH1* -93G>A polymorphism (rs1800734), as these genetic factors could play a causative role in *MLH1* promoter methylation. Whenever material was available, germline *MLH1* methylation status was studied and DNA sequencing for germline *GADD45A* mutations was performed. In order to analyze whether the younger patients exhibit an intrinsic higher methylation tendency in their genome, the methylation status of eight CIMP markers was determined in the tumors. In a selected subset of young patients, whole genome SNP array analysis was performed on formalin-fixed, paraffin-embedded (FFPE) tumor tissue and compared with previously published data to search for recurrent chromosomal aberrations involved in *MLH1* methylation.

Methods

Patient material

Tumor tissues were obtained from 46 sporadic right sided colon cancer patients analyzed between 1997 and 2006 at the Leiden University Medical Center (Leiden, The Netherlands). MSI analysis, additional MMR immunohistochemistry (IHC) and MMR germline mutation analysis were performed due to a relatively young age of onset and/or a suspected family history of Lynch syndrome. As we mainly focused on comparing relatively young patients with patients of older ages, our sample set was enriched for young patients with *MLH1* methylation. Our sample set contained a high percentage of *MLH1* methylated colon cancer patients with an age of onset below 50 years (28%, 13 cases) which is not a reflection the general age distribution of this of type colon cancer. The present study falls under approval by the Medical Ethical Committee of the LUMC (protocol P01-019). Informed consent was obtained according to protocols approved by the LUMC Medical Ethical Committee (02–2004). Patient samples were handled according to the medical ethics guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences (www.federa.org).

DNA isolation and MSI analysis

DNA was isolated from 0.6 mm FFPE punches after assessment of corresponding hematoxylin-eosin stained slides by a pathologist (HM). Standard deparaffination preceded DNA isolation using the Wizard Genomic DNA Purification kit (Promega, Madison, WI, US). The microsatellite instability status of each of the tumors was determined using the Promega MSI analysis system (Version 1.2, Promega, Madison, WI, US) following the recommendations of the National Cancer Institute/ICG-HNPCC[34-36]. Tumors with at least two out of five mononucleotide markers unstable were classified as MSI-H.

IHC of MMR proteins

Standard three-step, indirect IHC was performed on 4-µm tissue sections that had been transferred to glass slides, including citrate antigen retrieval, blockage of endogenous peroxidase and endogenous avidin-binding activity and di-aminobenzidine development. The following antibodies were used: anti-MLH1 (clone G168-728; BD Biosciences, San Jose, CA), anti-PMS2 (clone A16-4; BD Biosciences), anti-MSH2 (clone GB-12; 1:100; Oncogene Research Products, San Diego, CA) and anti-MSH6 (clone 44; 1:400; BD Biosciences). The utilized secondary antibodies were biotinylated rabbit anti-mouse IgG antibodies (DAKO, Glostrup, Denmark), goat anti-rabbit IgG antibodies (DAKO, Glostrup, Denmark) and biotinylated-peroxidase streptavidin complex (SABC; DAKO, Glostrup, Denmark). Loss of expression was assessed by a complete lack of staining in the tumor cell nuclei with concurrent staining in normal epithelium, stroma or infiltrating leukocytes.

Mutation analysis

BRAF V600E mutations were detected using flanking primers that have been previously described[37]. DNA sequence analysis of codons 12 and 13 of *KRAS* was performed as previously described[38]. For direct sequencing of *GADD45A*, six exon primer pairs were designed (encompassing 100 bp of intronic sequence) using the Primer3 web-tool for the amplification of the four exons of *GADD45A*[39]. The utilized primers are listed in Additional file 1. Primers utilized for sequence analysis of the *MLH1* -93G>A polymorphism were designed to amplify the region spanning from -231 bp to -51 bp from the *MLH1* transcription start site[39]. PCR products were purified with the QIAquick PCR Purification kit (Qiagen, Hilden, Germany). Sequencing was performed at the Leiden Genome Technology Center (LGTC, Leiden, The Netherlands) using an ABI 3730 XL (Applied Biosystems, Foster City, CA). Mutational analysis was performed using mutational surveyor (SoftGenetics LLC., State College, PA). Results of all mutational analyses are summarized in Table 1 (extended in Additional file 2).

Table 1: Occurrence of *BRAF* mutations, SNP rs1800734 in relation to age and *MLH1* methylation status.

		***MLH1* methylation**		***BRAF***	***MLH1* -93G>A**			
Type	**n**	**M**	**pM**	**Mut**	**G/G**	**G/A**	**A/A**	**NA**
Adenoma	2	2	0	1	0	1	1	0
Carcinoma	44	33	11	24	12	20	7	5
Age								
Total <50	13	9	4	4	4	7	2	0
Total ≥50	33	26	7	21	8	14	6	5
BRAF mut	25	23	2	25	7	12	4	2

Methylation analysis

Methylation of the 5' regulatory *MLH1* region at -200 bp (from the transcription start site) was analyzed by using Methylation-Specific PCR (MSP) primers that have been previously described[40]. Sample DNA (100 ng) was mixed with carrier DNA (salmon sperm DNA, 400 ng) followed by bisulfite conversion using the EZ DNA Methylation Gold kit (Zymo Research, Orange, US) and the standard protocol provided by the manufacturer. Amplified fragments were analyzed by electrophoresis through a 2% agarose gel and on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). The utilized primers are listed in Additional file 1.

Contamination of the carcinoma tissue by stromal or inflammatory cells was unavoidable in some cases, despite use of micro-dissection, and tumors with a partially methylated phenotype were scored as methylated.

Methylation of MINT1, MINT2, MINT12, MINT31, *RIZ1* and *TIMP3* was determined by MSP [14]. Primers and conditions are listed in Additional file 1. MINT27 and Megalin methylation was determined by Combined Bisulfite Restriction Analysis (COBRA)[41]. Tumors were determined to be CIMP-high when four or more markers besides *MLH1* showed methylation, and tumors were determined to be CIMP-low when containing three or fewer methylated markers besides *MLH1*. For validation of our CIMP marker set, methylation of *IGF2, SOCS1, NEUROG1, RUNX3, CACNA1G* was determined by MSP in the cases with an age of onset less than 50 years[42, 43].

SNP array analysis, copy number changes and loss of heterozygosity (LOH) assessment

Single nucleotide polymorphism array analysis, copy number change and loss of heterozygosity (LOH) assessment were performed as previously described[30]. For each sample, four SNP panels (linkage panel, LP), LP1-4, were tested. All LP panels were combined for testing of the entire genome. LP1 covers chromosomes 1 to 3 and 22, LP2 covers chromosomes 5 to 9, LP3 covers 10 to 15 and 21, and LP4 covers chromosomes 4, 16 to 20, X and Y. Each panel was separately analyzed on a bead array. Due to the limited availability of archival tumor tissue, some of the LPs could not be analyzed. In two cases two LPs and in one case one LP could not be analyzed. To assess the fraction of the genome altered, the number of chromosome cytobands that were altered was divided by the total number of cytobands tested.

Statistical analysis

Differences in mutation and MSI frequencies between groups were analyzed using Fisher's exact and Chi-Square tests. A p-value below 0.05 was considered to indicate statistical significance. Yates' correction was used whenever a value lower than 5 was used in the Chi-Square test.

Results

We characterized 13 MSI-H colon cancer cases with *MLH1* promoter methylation from patients with an age of onset below 50 years and compared these data with those obtained from 33 MSI-H cases of patients over 50 years of age. Based on the Bethesda guidelines, which recommend MSI testing for all colorectal cancers in patients diagnosed before 50 years of age, we used the cutoff age of 50 for our comparisons[36]. The mean age of the study cohort was 61 years (SD is 31 years). The majority of tumors originated from the proximal colon (n=36, 78% of total), while a low percentage (n=5, 11%) of MSI-H tumors originated distal from the splenic flexure. All tumors showed loss of expression of nuclear MLH1 and its heterodimer PMS2, confirming the deleterious effect of *MLH1* promoter methylation. Both MSH2 and MSH6 stained positive and pathogenic germline mutations in any of the four mismatch repair genes were identified in none of the patients.

Two patients identified with germline MLH1 epimutation

For seven MSI-H patients with an age of onset below 50 years and 13 patients aged above 50 years, normal colonic epithelium and/or leukocyte DNA was available for germline methylation analysis of *MLH1*. Two female patients were identified as having germline *MLH1* promoter methylation as both normal colonic epithelium and leukocyte DNA tested positive. The first patient (ID60, Figure 1) presented with a right sided colon cancer at the age of 33 and endometrial cancer at age 52. Her family history showed a sister with endometrial cancer at the age of 37. Apart from a maternal grandfather with colon cancer at the age of 90 years and a maternal niece with duodenum cancer at 39 years, no other tumors from within the Lynch syndrome spectrum were seen. The second patient with germline methylation (ID36) does not have a family history with characteristics of Lynch syndrome. She was diagnosed with colon cancer at age 60 (MSI-H with a *BRAF* V600E somatic mutation) and pancreatic cancer at age 62 (scored as MSS).

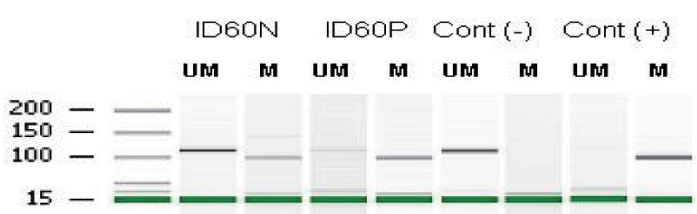


Figure 1 - Lab-on-chip results of a MLH1 MSP performed on normal tissue and peripheral blood from patient ID60. Lane one contains the lab-on-chip DNA marker. Partial methylation of both normal colon mucosa (ID60N) and peripheral blood (ID60P) was observed as they show products produced by the primer pairs amplifying unmethylated (UM) and methylated (M) template DNA. Negative (Neg) and Positive (Pos) controls represent unmethylated (Neg) and methylated (Pos) controls, respectively. The first lane is a visualization of the Agilent DNA 1000 Marker 15/1500 in base pairs (bp)

BRAF mutation shows an age-dependent trend in MSI-H tumors

Somatic *BRAF* V600E mutations were found in 25 out of the 44 tested tumors, whereas no *KRAS* codon 12/13 mutations were found (Table 1; extended information in Additional file 2). The majority of the *BRAF* mutations were identified in patients over 50 years of age (n=21, 65.6% of patients over 50). Comparison with the patient group under 50 years of age (n=4, 31% of patients under 50) showed a significant difference between the two groups (p-value: 0.044).

GADD45A somatic and germline DNA mutation analysis

GADD45A was successfully sequenced for 38 samples (17 normal epithelium samples and 21 tumor samples). Exon 1 was successfully studied for 37 samples and did not reveal any alterations. We identified 5 cases with an SNP (rs3783466, c.45-23C>T) present in the first intron. Exon 2 was studied for 38 samples and did not reveal any alterations. Exon 3 was sequenced for 38 samples and revealed a variant that was not previously described (in ID70). This heterozygous C>T transition resulted in a neutral amino acid change from proline to serine (p. Pro119Ser) and was predicted by the Sorting Intolerant From Tolerant (SIFT) prediction software to be a tolerated mutation[44]. Exon 4 was studied for 37 samples and did not reveal any alterations.

No association between age, *BRAF* and *GADD45A* mutation status or the *MLH1* -93G>A polymorphism was observed. An overview of the *GADD45A* mutation data is given in extended information in Additional file 2.

MLH1 -93G>A polymorphism analysis

We screened 41 out of 46 (13 below 50, 28 above) samples successfully for the *MLH1* -93G>A polymorphism (rs1800734), by sequence analysis. The G/G genotype was found in 12 samples (29.3%), G/A in 21 (51.2%) and the A/A genotype in 8 (19.5%). No significant differences between tumors grouped by age (with either 50 or 60 years as a cutoff for early onset colon cancer), *BRAF* or *GADD45A* mutational status were observed in our patients. An association between the G>A polymorphism and ages of onset above 50 years ($p=3.5x10^{-5}$) was found in comparison with published control samples (Table 2)[24]. This association was lost (p=0.19) when comparing younger patients with corresponding published control samples (Table 2)[24]. However, grouping of the A/A and G/A genotypes provided lower p-values when comparing both age groups to controls. A similar distribution of the A allele was found in the young age group as for the patients above 50 years. In this comparison the association between the A allele and the group with an age of onset below 50 years was significant (p=0.035), although the significance was lost after the required Yates' correction (p=0.068, Table 2). A numeric overview of the *MLH1* -93G>A polymorphism sequence data is given in Table 1 (extended information in Additional file 2 and Table 2).

Table 2: Genotype frequencies of MLH1 -93G>A polymorphism in sporadic MSI-H colon cancer with an age of onset below and above 50 years.

	n	Genotype frequency (%)			Chi-square	DF	P-value
		GG	GA	AA			
Age at diagnosis <50	13	4(31)	7(54)	2(15)			
Controls 1	929	554(59.5)	331(35.5)	44(5)	6.0	2	0.19427123*
Controls 2<60	501	287(57)	175(35)	39(8)	3.8	2	0.34061584*
Age at diagnosis ≥50	28	8(29)	14(50)	6(21)			
Controls 1	929	554(59.5)	331(35.5)	44(5)	20.5	2	0.00003502
Controls 2 >60	1462	883(60)	513(35)	66(5)	22.6	2	0.00001209
		GG	**GA+AA**		**Chi-square**	**DF**	**P-value**
Age at diagnosis <50	13	4(31)	9(69)				
Controls 1	929	554(60)	375(40)		4.4	1	0.06890141*
Controls 2<60	501	287(57)	214(43)		3.6	1	0.10499389*
Age at diagnosis ≥50	28	8(29)	20(71)				
Controls 1	929	554(60)	375(40)		10.8	1	0.00100409
Controls 2 >60	1462	883(60)	579(40)		11.6	1	0.00066844

n: Total number
Chi-square: Value of the chi-square test
DF: Degrees of freedom
Control samples are adapted from literature. Percentages are given in brackets. (Raptis et al., for controls 1 and Samowitz et al. for controls 2[36,38]). P-values with a * are Yates corrected.

CpG island DNA methylation is more frequent in older patients and is highly correlated with BRAF mutation in younger colon carcinoma patients

We examined the methylation status of 31 samples (11 below 50 years, 20 above) using six CIMP markers (MINT1, MINT2, MINT12, MINT31, *RIZ1* and *TIMP3*) with MSP and two CIMP markers (MINT27 and Megalin) with COBRA. Results of the analysis are presented in Table 3. Although all samples in this sub-selection of our MSI-H study group contain *MLH1* methylation, a clear age-related trend of methylation was observed. Out of the 11 tested patients that were below 50 years of age and had *MLH1* methylated colon cancer, only four were shown to be CIMP-high. Remarkably, all of these young CIMP-high cancers showed *BRAF* mutations, whereas such mutations were not detected in samples with less extensive methylation. A higher frequency of CIMP-high (20/20 vs. 4/11, $p=3.1x10^{-4}$ (Yates' corrected))

was observed for colon cancer patients above the age of 50, concomitant with the higher number of *BRAF* mutations found in these patients.
The methylation status of our early onset cases were validated by use of 5 additional CIMP markers (*IGF2, SOCS1, NEUROG1, RUNX3, CACNA1G*). All validated samples showed similar levels of methylation in both marker sets (Table 3). Although sample ID1 showed methylation of 3/5 additional markers, the cumulative amount of markers still led us to determine this sample as CIMP-low.

Genomic profiling of MSI-H colon carcinomas

For a sub-selection of 15 MSI-H carcinomas (5 below and 10 above 50 years of age) for which sufficient DNA was available, genome-wide profiles of copy number abnormalities and copy neutral LOHs (cnLOHs) were obtained using SNP arrays suitable for analysis of archival FFPE tissue (Additional file 3). Chromosomal copy number changes were observed in 7/15 samples. Physical chromosomal loss was a rare event (on average of 0.2% of the genome) and was only found in 3/15 carcinomas, in which small telomeric regions on chromosomes 1q, 4q, 8p and 18q were deleted. An overview of the events in all tested samples is given in Figure 2. Four chromosomal regions showed cnLOH in more than one tumor: chr 2q23.1-37.3 (n=2, ID50 and ID59), 3p21.31-26.3 (n=2; ID18 and ID39, containing *MLH1*), 9p21.2-24.3 (n=2; ID3 and ID36) and 11p15.1-15.5 (n=2 ; ID20 and ID59).

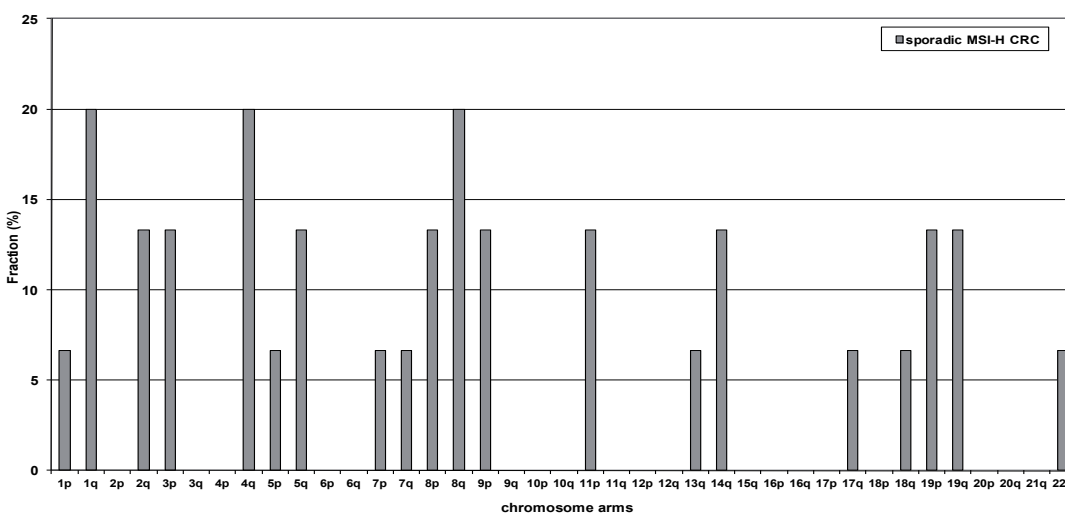


Figure 2: Chromosomal events per chromosome arm in 15 sporadic MSI-H carcinomas. Bars indicate the percentage of the tested sporadic MSI-H carcinomas containing a chromosomal aberration or copy neutral LOH per chromosomal arm.

Table 3: CIMP marker methylation.

ID	Gender	Age	Type	Location	*BRAF*	*KRAS*	*GADD45A* mutation	*MLH1* rs1800734	MLH1 methylation	MINT1	MINT2	MINT12	MINT31	*RIZ*	*TIMP3*	MINT27	*Megalin*	*IGF2*	*SOCS1*	*NEUROG1*	*RUNX3*	CACNA1G	CIMP
1	M	15	Ca	NA	wt	wt	C>T	G/G	pM														Lo
3	F	27	Ca	R	wt	wt	wt	G/A															Lo
7	M	36	Ca	L	wt	wt	C>T	A/A	pM														Lo
11	M	39	Ca	L	wt	wt	wt	G/A	pM														Lo
12	M	41	Ca	R	wt	wt	wt	G/G										NA	NA	NA	NA	NA	Lo
13	M	42	Ca	R	V600E	wt	C>T	G/A				NA											H
17	M	43	Ca	R	wt	wt	wt	G/G	pM														Lo
18	F	44	Ca	R	wt	wt	wt	G/A															Lo
20	F	46	Ca	R	V600E	wt	wt	G/A															H
21	M	47	Ca	R	V600E	wt	wt	G/A															H
23	F	48	Ca	R	V600E	wt	wt	G/G							NA								H
25	F	52	Ca	L	V600E	wt	wt	G/G															H
27	F	53	Ca	L	wt	wt	wt	G/G	pM														H
15	M	55	Ca	R	V600E	wt	wt	G/A	pM								NA						H
35	M	59	Ca	R	V600E	wt	wt	G/A															H
36	F	60	Ca	R	V600E	wt	C>T	G/A															H
37	F	60	Ca	R	wt	wt	wt	G/G															H
38	F	61	Ca	R	wt	wt	wt	G/A															H
39	M	62	Ca	R	V600E	wt	wt	G/G															H
42	F	62	Ca	NA	wt	wt	wt	G/A	pM														H
43	F	62	Ca	R	V600E	wt	wt	A/A								NA							H
65	F	64	Ca	R	V600E	wt	wt	G/G															H
47	M	67	Ca	NA	V600E	wt	wt	A/A															H
66	F	69	Ca	R	V600E	wt	wt	A/A															H
67	M	75	Ca	R	V600E	wt	wt	G/A															H
68	F	75	Ca	R	V600E	wt	wt	G/G	pM														H
69	F	76	Ca	R	wt	wt	wt	G/A															H
56	F	78	Ca	R	V600E	wt	NA	G/A	pM														H
57	F	80	Ca	R	wt	wt	wt	A/A															H
70	F	80	Ca	R	V600E	wt	C>T	G/G															H
59	M	84	Ca	R	V600E	wt	wt	G/A															H

NA: Not available
R: Right sided
L: Left sided
wt: Wildtype
M: Male
F: Female

CIMP: CpG island methylator phenotype
Lo: CIMP-low
H: CIMP-high
V600E: *BRAF* V600E mutation
C>T: rs3783466c.45-23C>T

Discussion

Since CpG island hypermethylation (including *MLH1*) in colon mucosa is considered to be age-related[9], the finding of hypermethylation of *MLH1* at a younger age is unexpected. Since 2002, several manuscripts pointed to the existence of *MLH1*germline methylation[15-21]. More recently, *MSH2* methylation due to an inherited deletion in the 3'end of *EPCAM/TACSTD1* was also discovered[4]. Methylation of *MLH1* can also be found in addition to a germline MMR mutation, as described by Rahner et al.[45]. We studied 13 MMR germline mutation-negative patients with MSI-H colon cancer (mostly right-sided) at ages of onset under 50 years. These data were compared with those obtained from a control group of 33 patients with an age of onset above 50. The presence of (somatic) promoter methylation of *MLH1* in the tumors made Lynch syndrome unlikely. We identified two female patients with ages of onset of 33 and 60 years harboring germline *MLH1* methylation. Relatively young patients without a strong family history who present a MSI-H tumor with loss of MLH1 and PMS2 protein expression are suggested as candidates for *MLH1* germline epimutation screening0[17, 21]. We identified one patient with germline *MLH1* methylation in seven tested cases who were less than 50 years of age, giving a frequency of ~14%. Although the low number of tested samples in this study makes this percentage not representative, this number is not significantly higher than the frequency range of 0.6-13% described in studies screening for germline *MLH1* methylation in Lynch syndrome-suspected patients[5]. The discovery of germline *MLH1* methylation in a patient aged 60 years at diagnosis is surprising, as the patients with germline *MLH1* methylation described prior to this study (n=25) have a mean age of diagnosis of 37 years with a range of 17-46[5].

In contrast to the group with an age of onset above 50 years, only some (4/11) of the *MLH1* methylated MSI-H tumors from patients below 50 years showed high levels of CIMP marker methylation (CIMP-high). For the patient group with an age of onset below 50 years the CIMP-high status completely overlapped with *BRAF* mutations. As both *BRAF* and *KRAS* mutations have been observed in the earliest identified colonic neoplasms, and recent papers have provided evidence that induction of the ras oncogenic pathway will result in DNA hypermethylation, a causative effect of *BRAF/KRAS* mutations is likely[24, 46-50]. Instead of widespread CpG island methylation in non-*BRAF* mutated tumors in the early onset patient group, methylation seems to be largely restricted to the *MLH1* locus. Although the existence of locus-restricted methylation may be a reflection of the Gaussian curve of methylation patterns in relation to age, this finding may suggest a distinct, non-BRAF associated mechanism of *MLH1* methylation. However, all tumors here were selected upon *MLH1* promoter methylation which may explain the fact that *MLH1* is methylated more frequent than all other CIMP genes. As the methylation mechanism is (at least partly) age related, and progressive, a similar selection of tumors methylated on one of the other CIMP markers would have also shown more frequent methylation on these than other CIMP markers including *MLH1* and occurring in tumors not reaching the CIMP-high classification yet. A progressive methylation and CIMP appearance according to age similar as that shown in Table 3 favors the argument that *MLH1* methylation in these young patients is a reflection of the Gaussian curve of methylation patterns in relation to age.

An alternative hypothesis concerning the association between *BRAF* mutation and DNA methylation is that promoter methylation and silencing of specific target genes such as *IGFBP7* by promoter methylation could favor the selection of activating *BRAF* mutations,

since the oncogenic effect of activated BRAF would be enhanced in the absence of IGFBP7's inhibitory function[51]. Since promoter hypermethylation is partly age related the occurrence of *IGFBP7* hypermethylation and *BRAF* mutation would also explain the diminished occurrence in the young sporadic MSI-H patient group[50]. This role of BRAF in aberrant methylation initiation will have to be elucidated in the future. The locus-specific, non-BRAF associated mechanism of *MLH1* methylation suggested in our study should be addressed in a larger group of early onset sporadic colon cancer patients with *MLH1* methylation to provide additional insights.

In patients of older ages, there is an association between somatic *MLH1* methylation and the *MLH1* -93G>A polymorphism[22, 24, 49]. Indeed, when we compared our group of patients above 50 years of age with the published control groups of Raptis et al., and Samowitz et al., we observed an enrichment of the A allele. We explored the possibility that the A allele was more prevalent in the sporadic MSI-H at early ages. However a similar distribution in both age groups was found, no significant enrichment could be found for the cases under 50 years. The hypothesis of Samowitz et al., which suggests an increased likeliness of *MLH1* methylation in the presence of a CIMP/*BRAF* mutation background and a *MLH1* -93 G>A polymorphism, excludes young onset patients because of low levels of *BRAF* mutations[24]. Although Samowitz did find a significant difference in A allele distribution between MSI-H colon cancer age groups, our cohort of sporadic MSI-H colon cancer patients with *MLH1* methylation excluded patients with a germline MMR gene mutation, which might explain the difference found between our studies.

Knockdown and overexpression experiments of *GADD45A* in Xenopus laevis led to the suggestion that deregulation of GADD45A's role in active DNA demethylation could give rise to aberrant methylation[25]. The absence of pathogenic somatic and germline mutations in human *GADD45A* observed in our study and data published during this study[26, 27] suggest that a role for *GADD45A* mutations in aberrant hypermethylation in human colon tumors is unlikely.

In a subset of tumors (including five with an age of onset under 50 years), whole genome SNP array analysis of FFPE tumor tissue was used to assess possible causative loci for *MLH1* methylation. Our copy number and cnLOH analysis identified patterns in agreement with literature describing limited chromosomal instability in sporadic MSI-H colon tumors with *MLH1* methylation. The extent of copy number abnormalities (CNA) identified here is in agreement with that found by Trautman et al. and by van Puijenbroek et al.[30, 33]. In patients under 50 years, no specific genomic pattern was identified, although two cases showed overlapping alterations at chromosome 4q. The smallest region of overlap (region 4q35.1-4q35.2) encompasses the cancer associated genes *TLR3*, *CDKN2AIP*, *ING2*, *CASP3* and *SORBS2*, none of which are thought to cause aberrant DNA methylation. The four regions of cnLOH that showed infrequent overlap in the 15 tumors tested are not known as such. The cnLOH of 3p21.31-26.3, found in a 44 and a 62 year old, encompasses the 3p22.2 region where *MLH1* is located. Such cnLOH is not typical for sporadic MSI-H colon carcinomas, but is more readily found in tumors containing pathogenic *MLH1* mutations[30]. We cannot rule out that the identified cnLOH regions may harbor loci involved in *MLH1* methylation. However, the odds are against such a suggestion.

Conclusion

Although our study did not identify a cause for *MLH1* methylation in sporadic MSI-H colon cancer with an age of onset below 50 years, we observed methylation to be almost restricted to the *MLH1* locus in patients without a *BRAF* mutation. We show that this early onset group consists of two sub-groups: those which are CIMP-high and contain a *BRAF* mutation (resembling sporadic MSI-H in the older age group to a great extend) and those with wild-type *BRAF* and limited methylation in addition to *MLH1* methylation.
Genomic analysis did not provide recurrent aberrations leading to identification of a possible cause of *MLH1* methylation in the cases under 50 years. Lastly, we excluded a role for somatic and germline *GADD45A* mutations in the tumorigenesis of early onset sporadic MSI-H colon cancer.

Competing interest

The authors declare that they have no competing interests.

Authors' contributions

EHJvR carried out the molecular genetic studies, the sequence alignment, statistical analysis and drafted the manuscript. MvP, DE and EJdM participated in the molecular genetic studies and the sequence alignment. AM and RvE carried out the hybridization of the SNP arrays. JO performed the statistical analysis of the SNP data. FH and CT performed the management of the clinical cases. KADW performed the MSP validation experiments on the additional CIMP markers, under supervision of MvE. TvW, JMB and HM conceived of the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Erica Borsi for validation of our *MLH1* methylation results in samples ID36, ID60 and ID61.

Supplementary data

Additional file 1: Primer table. All primers used in the current study listed. Available at: http://www.biomedcentral.com/1471-2407/10/180/additional/

Additional file 2: Overview of the cohort used. Gender, age, tumor location, *MLH1* methylation, *MLH1* rs1800734, BRAF mutation, KRAS mutation and GADD45A mutation status, when available, are given for each sample used. Available at: http://www.biomedcentral.com/1471-2407/10/180/additional/

Additional file 3: Table 1: Regions of copy number alterations and cnLOH.

ID	Gender	Age	*BRAF*	*GADD45A*	CIMP	**Gain**/*Loss*	cnLOH	Missing LP
3	F	27	wt	wt	L		9p21.2-9p24.3	3, 4
12	M	41	wt	wt	L			4
18	F	44	wt	wt	L	**chr:19**	3p21.31-3p26.3	
20	F	46	V600E	wt	H	**8q11.22-8q24.3** **1q21.2-1q32.2** *1q42.2-1q44*	1p36.12-1p36.33 4q35.1-4q35.2 11p15.1-11p15.5	
23	F	48	V600E	wt	H	**1q42.13-1q44** **4q31.21-4q31.23** *4q32.1-4q35.2* *18q23*	17q11.2-17q25.3	
25	F	52	V600E	wt	H	**chr:5 ; chr:8**		
36	F	60	V600E	C>T	H	**chr:7 ; chr:14**	9p13.2-9p24.3	
39	M	62	V600E	wt	H		3p21.32-3p26.3 14q31.1- 14q32.33	
42	F	62	wt	wt	H			
43	F	62	V600E	wt	H			
47	M	67	V600E	wt	H			
50	M	71	NA	wt	NA	**chr:19**	2q23.1-2q37.3 5q35.1-5q35.3 chr:13	
56	F	78	V600E	NA	H		5q21.3-5q22.3	3, 4
57	F	80	wt	wt	H	**8q24.21-8q24.3** *8p22-8p23.3*		
59	M	84	V600E	wt	H		2q14.3-2q37.3 11p12-11p15.5 chr:22	

F: Female
M: Male
L: CIMP-low
H: CIMP-high
wt: Wildtype
NA: Not available
LP: Linkage panel
V600E: *BRAF* V600E mutation
C>T: rs3783466c.45-23C>T

References

1. Aaltonen LA, Peltomaki P. Genes involved in hereditary nonpolyposis colorectal carcinoma. Anticancer Res 1994;14:1657-1660.
2. Akiyama Y, Sato H, Yamada T, Nagasaki H, Tsuchiya A, Abe R, Yuasa Y. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. Cancer Res 1997;57:3920-3923.
3. Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, . Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75-80.
4. Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, Lee TY, Bodmer D, Hoenselaar E, Hendriks-Cornelissen SJ, Tsui WY, Kong CK, Brunner HG, van Kessel AG, Yuen ST, van Krieken JH, Leung SY, Hoogerbrugge N. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nat Genet 2009;41:112-117.
5. Niessen RC, Hofstra RM, Westers H, Ligtenberg MJ, Kooi K, Jager PO, de Groote ML, Dijkhuizen T, Olderode-Berends MJ, Hollema H, Kleibeuker JH, Sijmons RH. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. Genes Chromosomes Cancer 2009;48:737-744.
6. Niv Y. Microsatellite instability and MLH1 promoter hypermethylation in colorectal cancer. World J Gastroenterol 2007;13:1767-1769.
7. Gervaz P, Bucher P, Morel P. Two colons-two cancers: paradigm shift and clinical implications. J Surg Oncol 2004;88:261-266.
8. Young J, Simms LA, Biden KG, Wynter C, Whitehall V, Karamatic R, George J, Goldblatt J, Walpole I, Robin SA, Borten MM, Stitz R, Searle J, McKeone D, Fraser L, Purdie DR, Podger K, Price R, Buttenshaw R, Walsh MD, Barker M, Leggett BA, Jass JR. Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. Am J Pathol 2001;159:2107-2116.
9. Kakar S, Burgart LJ, Thibodeau SN, Rabe KG, Petersen GM, Goldberg RM, Lindor NM. Frequency of loss of hMLH1 expression in colorectal carcinoma increases with advancing age. Cancer 2003;97:1421-1427.
10. Nakagawa H, Nuovo GJ, Zervos EE, Martin EW, Jr., Salovaara R, Aaltonen LA, de la CA. Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. Cancer Res 2001;61:6991-6995.
11. Shen L, Kondo Y, Hamilton SR, Rashid A, Issa JP. P14 methylation in human colon cancer is associated with microsatellite instability and wild-type p53. Gastroenterology 2003;124:626-633.
12. Toyota M, Ho C, Ahuja N, Jair KW, Li Q, Ohe-Toyota M, Baylin SB, Issa JP. Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. Cancer Res 1999;59:2307-2312.
13. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007;50:113-130.
14. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A 2007;104:18654-18659.
15. Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Res 2002;62:3925-3928.

16. Gylling A, Ridanpaa M, Vierimaa O, Aittomaki K, Avela K, Kaariainen H, Laivuori H, Poyhonen M, Sallinen SL, Wallgren-Pettersson C, Jarvinen HJ, Mecklin JP, Peltomaki P. Large genomic rearrangements and germline epimutations in Lynch syndrome. Int J Cancer 2008;124:2333-2340.
17. Hitchins M, Williams R, Cheong K, Halani N, Lin VA, Packham D, Ku S, Buckle A, Hawkins N, Burn J, Gallinger S, Goldblatt J, Kirk J, Tomlinson I, Scott R, Spigelman A, Suter C, Martin D, Suthers G, Ward R. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;129:1392-1399.
18. Hitchins MP, Wong JJ, Suthers G, Suter CM, Martin DI, Hawkins NJ, Ward RL. Inheritance of a cancer-associated MLH1 germ-line epimutation. N Engl J Med 2007;356:697-705.
19. Morak M, Schackert HK, Rahner N, Betz B, Ebert M, Walldorf C, Royer-Pokora B, Schulmann K, von Knebel-Doeberitz M, Dietmaier W, Keller G, Kerker B, Leitner G, Holinski-Feder E. Further evidence for heritability of an epimutation in one of 12 cases with MLH1 promoter methylation in blood cells clinically displaying HNPCC. Eur J Hum Genet 2008;16:804-811.
20. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004;36:497-501.
21. Valle L, Carbonell P, Fernandez V, Dotor AM, Sanz M, Benitez J, Urioste M. MLH1 germline epimutations in selected patients with early-onset non-polyposis colorectal cancer. Clin Genet 2007;71:232-237.
22. Raptis S, Mrkonjic M, Green RC, Pethe VV, Monga N, Chan YM, Daftary D, Dicks E, Younghusband BH, Parfrey PS, Gallinger SS, McLaughlin JR, Knight JA, Bapat B. MLH1 -93G>A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. J Natl Cancer Inst 2007;99:463-474.
23. Chen H, Taylor NP, Sotamaa KM, Mutch DG, Powell MA, Schmidt AP, Feng S, Hampel HL, de la CA, Goodfellow PJ. Evidence for heritable predisposition to epigenetic silencing of MLH1. Int J Cancer 2007;120:1684-1688.
24. Samowitz WS, Curtin K, Wolff RK, Albertsen H, Sweeney C, Caan BJ, Ulrich CM, Potter JD, Slattery ML. The MLH1 -93 G>A promoter polymorphism and genetic and epigenetic alterations in colon cancer. Genes Chromosomes Cancer 2008;47:835-844.
25. Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F, Niehrs C. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature 2007;445:671-675.
26. Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B, Liebermann DA. Conserved DNA methylation in Gadd45a(-/-) mice. Epigenetics 2009;4.
27. Jin SG, Guo C, Pfeifer GP. GADD45A does not promote DNA demethylation. PLoS Genet 2008;4:e1000013.
28. Carrier F, Georgel PT, Pourquier P, Blake M, Kontny HU, Antinore MJ, Gariboldi M, Myers TG, Weinstein JN, Pommier Y, Fornace AJ, Jr. Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin. Mol Cell Biol 1999;19:1673-1685.
29. Hollander MC, Fornace AJ, Jr. Genomic instability, centrosome amplification, cell cycle checkpoints and Gadd45a. Oncogene 2002;21:6228-6233.
30. van Puijenbroek M, Middeldorp A, Tops CM, van ER, van der Klift HM, Vasen HF, Wijnen JT, Hes FJ, Oosting J, van WT, Morreau H. Genome-wide copy neutral LOH is infrequent in familial and sporadic microsatellite unstable carcinomas. Fam Cancer 2008.
31. Douglas EJ, Fiegler H, Rowan A, Halford S, Bicknell DC, Bodmer W, Tomlinson IP, Carter NP. Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. Cancer Res 2004;64:4817-4825.
32. Li LS, Kim NG, Kim SH, Park C, Kim H, Kang HJ, Koh KH, Kim SN, Kim WH, Kim NK, Kim H. Chromosomal imbalances in the colorectal carcinomas with microsatellite instability. Am J Pathol 2003;163:1429-1436.

33. Trautmann K, Terdiman JP, French AJ, Roydasgupta R, Sein N, Kakar S, Fridlyand J, Snijders AM, Albertson DG, Thibodeau SN, Waldman FM. Chromosomal instability in microsatellite-unstable and stable colon cancer. Clin Cancer Res 2006;12:6379-6385.
34. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248-5257.
35. Bacher JW, Flanagan LA, Smalley RL, Nassif NA, Burgart LJ, Halberg RB, Megid WM, Thibodeau SN. Development of a fluorescent multiplex assay for detection of MSI-High tumors. Dis Markers 2004;20:237-250.
36. Umar A, Boland CR, Terdiman JP, Syngal S, de la CA, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261-268.
37. Xu X, Quiros RM, Gattuso P, Ain KB, Prinz RA. High prevalence of BRAF gene mutation in papillary thyroid carcinomas and thyroid tumor cell lines. Cancer Res 2003;63:4561-4567.
38. van Puijenbroek M, Nielsen M, Tops CM, Halfwerk H, Vasen HF, Weiss MM, van WT, Hes FJ, Morreau H. Identification of patients with (atypical) MUTYH-associated polyposis by KRAS2 c.34G > T prescreening followed by MUTYH hotspot analysis in formalin-fixed paraffin-embedded tissue. Clin Cancer Res 2008;14:139-142.
39. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-386.
40. Deng G, Chen A, Hong J, Chae HS, Kim YS. Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. Cancer Res 1999;59:2029-2033.
41. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997;25:2532-2534.
42. Ahlquist T, Lind GE, Costa VL, Meling GI, Vatn M, Hoff GS, Rognum TO, Skotheim RI, Thiis-Evensen E, Lothe RA. Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers 23. Mol Cancer 2008;7:94.
43. Hughes LA, van den Brandt PA, de Bruine AP, Wouters KA, Hulsmans S, Spiertz A, Goldbohm RA, de Goeij AF, Herman JG, Weijenberg MP, van EM. Early life exposure to famine and colorectal cancer risk: a role for epigenetic mechanisms 24. PLoS One 2009;4:e7951.
44. Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. Genome Res 2002;12:436-446.
45. Rahner N, Friedrichs N, Steinke V, Aretz S, Friedl W, Buettner R, Mangold E, Propping P, Walldorf C. Coexisting somatic promoter hypermethylation and pathogenic MLH1 germline mutation in Lynch syndrome. J Pathol 2008;214:10-16.
46. O'Brien MJ. Hyperplastic and serrated polyps of the colorectum. Gastroenterol Clin North Am 2007;36:947-68, viii.
47. Ordway JM, Williams K, Curran T. Transcription repression in oncogenic transformation: common targets of epigenetic repression in cells transformed by Fos, Ras or Dnmt1. Oncogene 2004;23:3737-3748.
48. Gazin C, Wajapeyee N, Gobeil S, Virbasius CM, Green MR. An elaborate pathway required for Ras-mediated epigenetic silencing. Nature 2007;449:1073-1077.
49. Nagasaka T, Koi M, Kloor M, Gebert J, Vilkin A, Nishida N, Shin SK, Sasamoto H, Tanaka N, Matsubara N, Boland CR, Goel A. Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950-60, 1960.

50. Velho S, Moutinho C, Cirnes L, Albuquerque C, Hamelin R, Schmitt F, Carneiro F, Oliveira C, Seruca R. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis?
25. BMC Cancer 2008;8:255.
51. Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, Whitehall VL, Leggett BA, Laird PW. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PLoS One 2009;4:e8357.

Chapter 4

BRAF mutation-specific promoter methylation of *FOX* genes

manuscript in preparation

Eddy H. J. van Roon[1,2,], Robert E. Ernst, Tom van Wezel[2], Hans Morreau[2], Judith M. Boer[1,4, §]

[1]Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
[2]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Pediatric Oncology, Erasmus MC, Rotterdam, The Netherlands [4]Netherlands Bioinformatics Center [§]Corresponding author

BRAF MUTATION-SPECIFIC PROMOTER METHYLATION OF *FOX* GENES

Abstract

Cancer-specific hypermethylation of (promoter) CpG islands is common during the tumorigenesis of colon cancer. Although associations between certain genetic aberrations, such as *BRAF* mutation and microsatellite instability, and the CpG island methylator phenotype (CIMP) have been found, the mechanisms by which these associations are established are still unclear.

Using differential methylation hybridization on oligonucleotide microarrays, we generated methylation profiles of paired tumor and normal colon. The majority of CpG regions found differentially methylated between *BRAF* mutant and wildtype tumors showed hypermethylation in the mutant cases. Enrichment of several cancer-related pathways, including the PI3 kinase and Wnt signaling pathways, was found. To focus on genes that are silenced in a tumor-specific rather than a lineage-specific manner, we used information on an epigenetic silencing mark in embryonic stem (ES) cells. Among the genes showing *BRAF* mutation-specific promoter methylation but no H3K27^{me3} mark in ES cells were forkhead box (FOX) transcription factors associated with the PI3-kinase pathway as well as *MLH1*, and *SMO*. Epigenetic down-regulation of these targets may contribute to mutationally active BRAF-driven tumorigenesis, explaining its association with aberrant DNA methylation.

Introduction

The CpG island methylator phenotype (CIMP) was introduced in 1999 by Toyota *et al*. to describe a subset of colorectal tumors with high levels of cancer-specific methylation[1]. Subsequent studies regarding (CIMP) in colon cancer described a strong association between this epigenetic phenotype, *BRAF* mutations and microsatellite instability (MSI)[2-8]. As sporadic MSI colon cancer is caused by promoter methylation of a mismatch repair gene (*MLH1*, *MSH2* or *MSH6*) the association between MSI and the high levels of DNA methylation in CIMP is considered a causative one[9, 10]. However, the association between activating *BRAF* mutations and CIMP remains unclear.

The field of epigenetic research has progressed from a candidate-gene to a genome-wide approach which not only provided a plethora of new candidate targets of cancer-specific DNA methylation but a better understanding of transcription regulation by DNA methylation as well.[11]. Using such genome-wide DNA methylation approaches could help identify new targets of *BRAF* mutation-specific promoter methylation. Hinoue *et al*. examined the CIMP- and *BRAF* mutation-specific methylation status of 1,505 CpG sites, located at 807 genes, in 235 primary colorectal tumors and discovered specific methylation of genes mediating various signaling pathways involved in colon cancer tumorigenesis[2]. In this study, we screened 32,171 CpG sites located at 10,537 genes in 19 colon cancer patients to obtain additional insight into the association between *BRAF* mutations and DNA methylation in colon cancer tumorigenesis.

Recent publications have reported a possible pre-marking of cancer-specific hypermethylated genes by the inactivation mark Histone H3 lysine 27 trimethylation (H3K27^{me3}) and binding

of polycomb group member SUZ12 in both ES cells and differentiated normal colon mucosal tissue[12-14]. These studies led to the suggestion that colon cancer cells utilize a pre-existing repression program to target loci for cancer specific promoter methylation[12, 14-16]. However, the presence of such repressive histone modifications at promoters during differentiation from ES to normal colon epithelium suggests that the associated genes are at a transcriptional silent state prior to tumor formation, reducing the relevance of the DNA methylation of pre-marked genes on tumorigenesis. In an attempt to identify biologically relevant *BRAF* mutation-specific promoter methylation, we excluded loci with H3K27^{me3} pre-marking in ES cells from the functional pathway analyses. By both extending the number of screened loci and filtering out pre-marked genes we identified new targets of *BRAF* mutation-specific methylation that could either create a favorable setting for the acquisition of *BRAF* mutations or function as an addition to up-regulation of the RAS-RAF-MEK pathway.

Materials and methods

Patient material

Anonymized tumor and normal fresh-frozen colon mucosa samples were obtained from patients who underwent surgery between 1988 and 2006 at the Leiden University Medical Center (Leiden, The Netherlands) or at the Rijnland Hospital (Leiderdorp, The Netherlands). Age, gender, histology, microsatellite instability (MSI), and *BRAF* V600E status for the 19 patients used for the array profiling are listed in Supplementary Table S1. Prior to DNA isolation, frozen sections were micro-dissected to minimize the presence of normal epithelium and stromal cells. To correct for age-dependent methylation, we used normal mucosa, distant from the tumor, from the same individuals. DNA was isolated by phenol/chloroform extraction and ethanol precipitation from 10-20 sections of 30 μm and yielded 10-50 μg. The present study was approved by the Medical Ethics committee of the LUMC (protocol P01-019). Cases were analyzed following the medical ethical guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences.

BRAF mutation analysis

BRAF V600E mutations were detected using flanking primers that have been previously described[17]. PCR products were purified with the QIAquick PCR Purification kit (Qiagen, Hilden, Germany). Sequencing was performed at the Leiden Genome Technology Center (LGTC, Leiden, The Netherlands) using an ABI 3730 XL (Applied Biosystems, Foster City, CA). Mutational analysis was performed using mutational surveyor (SoftGenetics LLC., State College, PA). Results are summarized in Supplementary Table S1.

Array hybridization

Differential methylation hybridization (DMH) was performed according to Yan *et al.* [18]. DNA (500 ng) was digested with *Mse*I, ligated to linkers, and sequentially digested with two methylation-sensitive restriction enzymes (HpaII and BstUI, New England Biolabs, Ipswich, MA, USA). Digested linker-ligated DNA was used as template for polymerase chain reaction (PCR) amplification (20 cycles) and coupled to fluorescent dyes. Cy5- or Cy3-labeled amplicons, representing methylated DNA fragments derived from tumor and

normal samples, were co-hybridized to the Agilent 244k human CpG island microarrays (Agilent Technologies, Santa Clara, CA, USA) in a dye-swap setup. Detection was done on a G2565BA scanner (Agilent Technologies) and feature extraction using Feature Extraction Software version 9.5.3.1 (Agilent Technologies).

Array data analysis

Non-background corrected data was preprocessed by within-array loess normalization followed by between-array aquantile normalization using limma v3.2.1[19] in R2.10.0[20]. Data was corrected for gene-specific dye bias using R package dyebias v1.4.0[21]. Raw data and preprocessed log2 ratios (tumor versus normal) per probe are available via GEO under accession number GSExxxxx. Probes mapping to the same *Mse*I fragment were expected to show similar hybridization patterns and not be independent. Therefore, we mapped probes to the human genome (UCSC assembly March 2006) cut *in silico* with *Mse*I. Fragments of 150-3000 bp mapping at least one complete probe and containing at least one BstUI or HpaII restriction site (n=32,171) were selected. In total, 195,625 of the 244,000 array probes (80.2%) mapped to such informative fragments, mostly with 1-2 probes per fragment, up to 33. For statistical analysis and visualization, the median log ratio per fragment was used to represent the fragment. Methylation differences between tumor and normal samples and tumor subgroups were analyzed using a linear model in limma v3.2.1[19]. The obtained P-values were corrected for multiple-testing[22] and fragments with a false discovery rate (FDR) <= 0.01 were selected as significant.

MLH1 and CIMP marker methylation

DNA samples (500 ng) were converted using the EZ DNA methylation Gold bisulphite kit (Zymo Research). For validation of methylation changes we performed a methylation-specific PCR (MSP) on the *MLH1* promoter using primers previously described[23] (Supplementary Table S1). Methylation of previously described CIMP markers:MINT1, MINT2, MINT12, MINT31, *RIZ1* and *TIMP3* was determined by MSP, while MINT27 and Megalin methylation was determined by Combined Bisulfite Restriction Analysis (COBRA)[6, 24]. Amplifications were carried out in a DNA Engine Dyad Peltier Thermal Cycler (Bio-Rad, Hercules, CA, USA) using AmpliTaq Gold PCR buffer and enzyme (Applied Biosystems, Foster City, USA). Amplified bands were visualized on a 2% agarose gel.

Exploratory data analysis

Differentially methylated fragments were compared to publicly available data containing chromosomal regions identified in chromatin immunoprecipitation using antibodies against $H3K27^{me3}$, $H3K4^{me3}$, CTCF and SUZ12 in ES cells followed by high-throughput sequencing[25-27]. By using the sqldf R package (version 0.3-5), we determined overlap of at least 20 bp between CpG fragments represented on the array and these regions. Enrichment of chromatin domains among the differentially methylated fragments was calculated by χ-squared test. Functional annotation clustering was performed in Panther 6.0[28]. Filtering of the differential methylation datasets by $H3K27^{me3}$ in ES cells using the dataset from Zhao *et al*.[27] was performed in R using the sqldf package.

Results

Colon cancer-specific CpG island methylation

We identified 1770 CpG-rich fragments with significant methylation differences between tumor and paired normal colon. Of these, 1234 fragments were associated with 816 genes, of which 531 were localized to gene promoters (Supplementary Table S2). As expected, CpG islands were mostly found hypermethylated in tumors (78.8%)[11].

We compared our results with those of Irizarry et al[11] who described 2707 cancer-specific differentially methylated regions (cDMRs) based on the comparison of 13 colorectal cancer tumor-normal pairs. Of the described cDMRs, 1203 overlapped with our CpG island array fragments, of which 282 (23%) were also differentially methylated between tumor and normal in our analysis. This overlap is reasonable considering the different, modest sized patient groups, and different experimental approaches.

CIMP-specific methylation

Next, we compared the tumor-normal methylation ratios between different groups of patients. Between CIMP-positive (n=11) and CIMP-negative (n=8) patients 749 CpG-rich fragments showed methylation changes, of which 85.6% had a higher tumor-normal methylation ratio in the CIMP-positive group. Of these fragments, 589 were associated with 508 genes, of which 244 were localized to gene promoters (Supplementary Table S3). In 8 out of 11 CIMP-positive tumors, promoter methylation of *MLH1*, the cause of microsatellite instability (MSI) in sporadic colon cancer, was observed which was consistent with methylation-specific PCR (Supplementary Table S1). We conclude that the hypermethylation in specific genomic regions used to define CIMP[6] is associated with methylation changes throughout the genome.

BRAF mutation-specific methylation

Activating *BRAF* mutations have been associated with high levels of CpG island methylation and MSI in colon cancer[2-8]. To investigate this association we compared the tumor-normal methylation ratio profiles of *BRAF* wildtypes (n=11) with those containing the *BRAF* V600E mutation (n=8). We identified 758 fragments with a *BRAF* mutation-specific methylation change, of which 96.3% had a higher tumor-normal methylation ratio in the *BRAF* mutant group. Out of these 758 fragments 579 were associated with 479 genes, of which 229 were localized to gene promoters (Supplementary Table S4).

Since *BRAF* mutations and CIMP co-occurred in eight samples, there was a high level of overlap between CIMP- and *BRAF* mutation-specific methylation changes (Figure 1A). Comparable levels of overlap were found focusing on promoter regions only (data not shown).

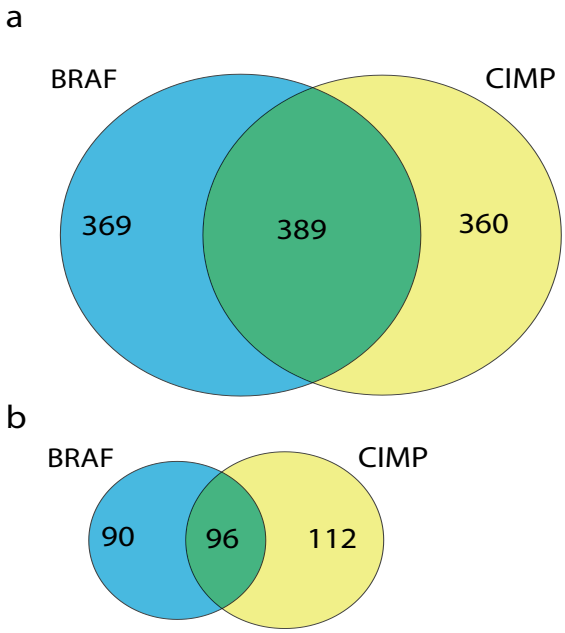


Figure 1- Proportional Venn diagrams showing the overlap between *BRAF* mutation- (blue) and CIMP-specific (yellow) methylation changes for all CpG-rich fragments (**A)** and promoter fragments filtered for H3K27^{me3} binding in ES cells (**B**).

Regions with tumor, CIMP and BRAF mutation-associated methylation changes are enriched for SUZ12 and H3K27^{me3} while depleted for CTCF and H3K4^{me3}

Regions binding the polycomb repressor complex 2 (PRC2) component SUZ12 in ES cells[29] were found to be enriched among the loci differentially methylated between colon cancer and normal colon (Table 1). The histone mark H3K27^{me3} is mediated by the PRC2 complex[29], and the two marks have been reported to be highly correlated[13]. Enrichment of ES cell H3K27^{me3} binding regions among the fragments with colon cancer-associated methylation changes was therefore expected and indeed observed. Similarly, fragments with CIMP- and *BRAF* mutation-associated differential methylation changes were also highly enriched for regions binding SUZ12 and H3K27^{me3} in ES cells (Table 1). Additionally, sites binding CTCF and the active chromatin mark H3K4^{me3} were underrepresented among the differentially methylated fragments. Interestingly, although all colon cancer-, CIMP- and *BRAF* mutation-specific fragments are underrepresented for H3K4^{me3}, this depletion is most evident for *BRAF* mutation-specific fragments.

After exclusion of fragments with H3K27^{me3} pre-marking in ES cells the overlap between CIMP- and *BRAF* mutation-specific methylation changes for all loci (Figure 1A) and promoters (Figure 1B) remained highly significant. Despite this high level of overlap approximately 50% of *BRAF* mutation-specific methylation changes showed no overlap with CIMP. In our functional analysis we focused on all promoter fragments with *BRAF* mutation-specific methylation changes regardless of overlap with CIMP.

	Total	SUZ12	p-val*	H3K27me3	p-val	CTCF	p-val	H3K4me3	p-val
Colon tumor fragments	1770	540(30.5%)	0	561(31.7%)	0	379(21.4%)	0	711(40.2%)	0.00001285
Remaining fragments	30401	3447(11.3%)		3336(11%)		10137(33.3%)		13835(45.5%)	
CIMP-specific fragments	749	160(21.4%)	0	169(22.6%)	0	183(24.4%)	0.0000011	300(40.1%)	0.0045882
Remaining fragments	31422	3827(12.2%)		3728(11.9%)		10333(32.9%)		14246(45.3%)	
BRAF-specific fragments	758	171(22.6%)	0	180(23.7%)	0	195(25.7%)	0.00003543	169(22.3%)	0
Remaining fragments	31413	3816(12.1%)		3717(11.8%)		10321(32.8%)		14377(45.8%)	

* p-val: P-values of the χ-square test

Table 1- Number of fragments with colon tumor-, CIMP- and *BRAF* mutation-specific methylation changes containing overlap with indicated chromatin marks in ES cells.

BRAF mutation-specific promoters (125)	Associated genes	Exp hits‡	Hits	p-val†
Hedgehog signaling pathway	*SMO ; GSK3A* ; CREBBP**	0.16	3	5.61E-04
PI3-kinase pathway	*FOXB1 ; FOXB2 ; FOXD3 ; CCND1 ; GSK3A**	0.72	5	8.49E-04
Insulin/IGF pathway-protein kinase B signaling Cascade	*FOXB1 ; FOXB2 ; FOXD3 ; GSK3A**	0.56	4	2.52E-03
Wnt signaling pathway	*NKD2* ; GNG4 ; CCND1 ; GSK3A* ; CREBBP* ; AXIN1 ; LEF1*	1.99	7	4.05E-03
Transcription regulation by bZIP transcription factor	*MTERF ; CREBBP* ; TAF7*	0.33	3	4.71E-03

*Asterisks mark the targets with low log2 in the *BRAF* wildtype group, suggesting that additional mechanisms such as genomic loss may play a role

† p-val: P-values of the binomial test between the *BRAF* mutation-specific gene list and the reference gene list

‡Exp hits: Expected number of hits by chance in the reference gene list

Table 2 - Pathways enriched with genes showing *BRAF* mutation-associated promoter methylation after exclusion of ES-cell H3K27^{me3} binding promoter fragments.

BRAF mutation-associated methylation pathway analysis

To identify biological pathways affected by *BRAF* mutation-associated gene methylation, we used 186 promoter regions that did not bind H3K27^{me3} in ES cells representing 125 genes after exclusion of duplicates and annotation by Panther 6.0. We found five significantly enriched pathways (P-val < 0.01, Table 2) containing 13 unique genes (Table 2).

With seven genes, the WNT pathway contained the most *BRAF* mutation-specific methylation changes (Table 2). However, the tumor-normal log2 ratios (Figure 2) of *AXIN1*, *CREBBP*, *GSK3A* and *NKD2* in the *BRAF* wildtype samples were low (-0.26 median, 0.12 standard deviation) compared to those in the *BRAF* mutated samples (-0.02 median, 0.12 standard deviation). While this could indicate tumor hypomethylation in *BRAF* wildtype samples compared to normal and *BRAF* mutated samples, the high level of chromosomal instability among *BRAF* wildtype samples suggests that copy-number loss is the most plausible explanation for this. To filter for this phenomenon we excluded fragments with a log2 ratio below one standard deviation of the median log2 ratio of all *BRAF* mutation-specific fragments in the *BRAF* wildtype group. A significant increase in the *BRAF* mutant log2 ratios compared to those of the *BRAF* wildtypes (approximately 0), indicate *BRAF* mutation-specific hypermethylation in these colon cancer samples (Figure 2). Upon filtering nine of the pathway associated genes remained (*SMO, FOXB1, FOXB2, FOXD3, CCND1, GNG4, LEF1, MTERF, TAF7*) and the PI3 kinase pathway was the only statistically significant enriched (P-val: 5.5E-03) pathway. Interestingly, besides promoter methylation of PI3 kinase pathway-associated forkhead box (FOX) genes we identified promoter methylation of three other FOX transcription factors: *FOXA1*, *FOXC1* and *FOXF1*. However, these promoters were bound by H3K27^{me3} and were excluded from our pathway analysis.

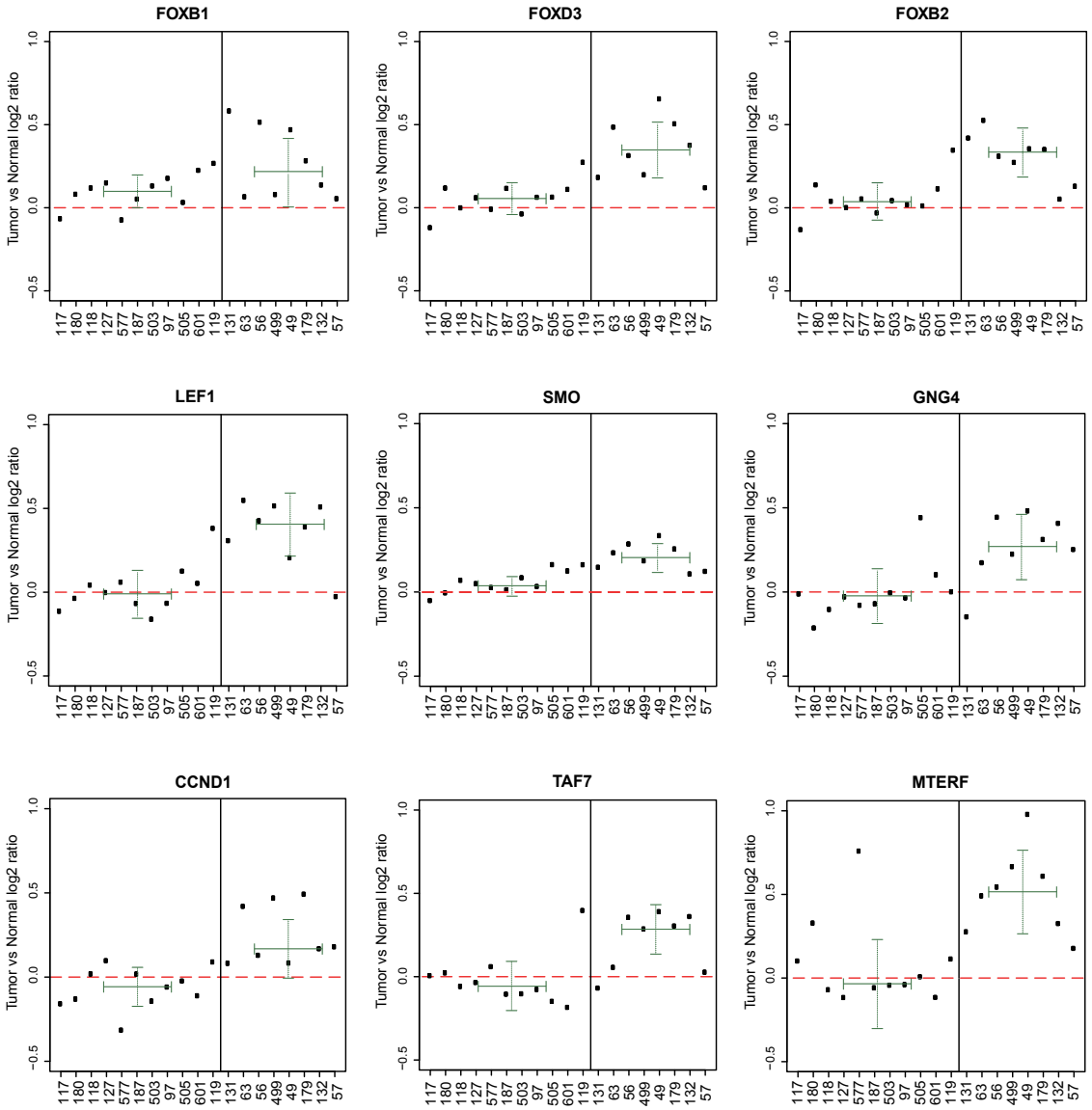


Figure 2 - Scatter plots for 9 unique genes with *BRAF* mutation-specific promoter methylation causing pathway enrichment. The Y-axis represents the tumor-normal log2 ratio for the median probe per fragment. Sample IDs are given below the X-axis with *BRAF* wildtypes left of the black line and *BRAF* mutants on the right. Median log2 ratios and standard deviations (dotted lines) for the *BRAF* wt and *BRAF* mutant groups are given in dark green.

Discussion

In this study we extended the number of screened CpG loci compared with previous studies performed in context of *BRAF* mutations to identify new *BRAF* mutation-specific methylation changes. This association between DNA methylation and activating *BRAF* mutations in colon cancer has been identified in multiple studies[2-8]. Here, we attempted to identify additional targets of *BRAF* mutation-specific DNA methylation that could provide a favorable context to either obtain a *BRAF* mutation or to attain the full potential of RAS-RAF-MEK induced proliferation provided by this activating mutation. Identified targets of promoter methylation showing pre-marking by H3K27^{me3} in ES cells were excluded to filter out methylation changes with minimal expected effects on transcription and thereby tumorigenesis[13]. We showed high levels of overlap between CIMP- and *BRAF* mutation-specific methylation changes which remained after filtering out pre-marked loci. Although Rada-Iglesias *et al.* showed a higher pre-marking of colon cancer-specific DNA methylation by H3K27^{me3} binding in normal colon epithelium compared to ES cells, we were restricted to using ES cell data due to incompatibility between data formats in our analyses[13]. Interestingly the promoter region of *MLH1*, found methylated in both a CIMP- and *BRAF* mutation-specific manner, was not filtered out. Therefore, *MLH1* promoter methylation, the cause of sporadic MSI colon cancer, is not established through utilization of a pre-existing repressive program in ES cells.

The study by Hinoue *et al.*[2] described *BRAF* mutation-specific DNA methylation of 60 genes in a comparison of 1505 CpG sites between 33 *BRAF* mutated tumors and 202 *BRAF* wildtype tumors. The identification of promoter methylation of the mediator of BRAFV600E-induced senescence, *IGFBP7*, led them to suggest that this epigenetic silencing provides a favorable context for the acquisition of *BRAF* mutations[2, 30]. Despite differences in experimental techniques and coverage, ten genes overlapped with our set of *BRAF* mutation-specific methylated fragments, including the RAS-RAF hyperactivation-associated *BMP3*, receptor kinases *EPHA3* and *FLT3* as well as the Hedgehog (Hh) signaling protein *SMO*. However, no overlap was found for the mediator of RAS-RAF oncogene-induced senescence, *IGFBP7*, possibly due to lack of *IGFBP7* promoter area coverage in our assay. Additionally, *BMP3* and *EPHA3* were pre-marked by H3K27^{me3} in our analysis suggesting minimal impact on gene expression and tumorigenesis.

We initially identified enrichment of five cancer-associated pathways by *BRAF* mutation-specific promoter methylation of 9 unique genes. Our analysis took into account copy number changes and filtered for this which could improve the reproducibility of DMH based assays[18, 31]. Exclusion of these loci resulted in the PI3-kinase pathway as the only pathway enriched in our analysis.

Among the four genes enriched in this pathway are the FOX transcription factors *FOXD3*, *FOXB1* and *FOXB2*. A recent study described FOXD3 as a p53 and p21^{cip1}-dependent negative cell cycle regulator which is suppressed by activated BRAF in melanoma cells[32]. Down-regulation of FOXD3 levels by promoter methylation could provide a favorable setting for either acquisition of a *BRAF* mutation or proliferation by RAS-RAF-MEK over-activation in colon cancer, similar to *IGFBP7*[2]. Interestingly, the FOXO transcription factors have been described as mediators of p21^{cip}-dependent BRAF induced senescence as well, indicating that multiple FOX genes are involved in this process[33]. We identified additional FOX genes with *BRAF* mutation specific promoter methylation that were not annotated as part of

the PI3-kinase pathway in our analysis as they were pre-marked by H3K27^{me3} in ES cells: *FOXA1*, *FOXC1* and *FOXF1*. However, the promoters of these genes were also pre-marked with H3K4me^{3} indicating possible tissue-specific expression of these genes. All three are targets of inactivation in breast cancer and both *FOXC1* and *FOXF1* subjected to promoter methylation[34-36]. Most intriguing is the description of FOXF1 as an inducer of G1-S and S-G2 cell cycle arrest, indicating a possible role in oncogene induced senescence[35]. Additional research is required to determine the role of these FOX genes in colon cancer-associated oncogene induced senescence and what the impact of their promoter methylation is on this mechanism. Finally, research into the sequence of such events is required to provide a better insight in the association between activating *BRAF* mutations and DNA methylation in colon cancer.

Acknowledgements

The authors would like to thank Michiel van Galen (Leiden Genome Technology Center Leiden, The Netherlands) for support with mapping fragments, and Philip Lijnzaad (Department of Physiological Chemistry, University Medical Center Utrecht, Utrecht, The Netherlands) for support with the dyebias R package.

Supplementary data

Supplementary Table S1 - Sample information. Gender (indicated by M for male and F for female), Age, Histology (indicated by Ad for adenoma and Ca for carcinoma), *MLH1* promoter methylation status and *BRAF* mutation status is given.

Patient ID	Gender	Age	Histology	*MLH1* methylation	*CIMP* status	*BRAF* V600E
117	M	69	Ad	U	-	wt
180	F	86	Ad	U	-	wt
118	F	69	Ad	U	-	wt
127	F	87	Ad	U	-	wt
577	M	32	Ad	U	-	wt
187	F	61	Ca	U	-	wt
503	F	63	Ca	U	-	wt
97	F	68	Ca	U	-	wt
131	F	71	Ca	U	+	mut
505	F	83	Ca	U	+	wt
63	F	57	Ca	U	+	mut
57	F	75	Ca	pM	+	mut
601	F	83	Ad	M	+	wt
56	M	75	Ca	M	+	mut
499	F	69	Ca	M	+	mut
119	F	76	Ca	M	+	wt
49	F	80	Ca	M	+	mut
179	M	90	Ca	M	+	mut
132	F	64	Ca	M	+	mut

Supplementary Table S2 - **Fragments with tumor-specific methylation changes**. Genomic location is provided by chromosome number (Chr) and the start (start) and end (end) position in the human genome build 18 (UCSC assembly: March 2006) basepair number. Annotation as promoter, divergent promoter or otherwise and the associated gene (UniGene and EntrezGene numbers) in case of promoter or intergenic (annotated as inside) is given as well as fragments length (length), number of DNA methylation-specific restriction sites (BstUI and HpaII), number of probes located in the specific fragment (Probes) and pre-marking by H3K27me^{3} in ES cells. Available upon request

Supplementary Table S3 - Fragments with CIMP-specific methylation changes. Genomic location is provided by chromosome number (Chr) and the start (start) and end (end) position in the human genome build 18 (UCSC assembly: March 2006) base pair number. Annotation as promoter, divergent promoter or otherwise and the associated gene (UniGene and EntrezGene numbers) in case of promoter or intergenic (annotated as inside) is given as well as fragments length (length), number of DNA methylation-specific restriction sites (BstUI and HpaII), number of probes located in the specific fragment (Probes) and pre-marking by H3K27me^3 in ES cells. Available upon request

Supplementary Table S4 - Fragments with BRAF mutation-specific methylation changes. Genomic location is provided by chromosome number (Chr) and the start (start) and end (end) position in the human genome build 18 (UCSC assembly: March 2006) base pair number. Annotation as promoter, divergent promoter or otherwise and the associated gene (UniGene and EntrezGene numbers) is given as well as fragments lengths, number of DNA methylation-specific restriction sites (BstUI and HpaII), number of probes located in the specific fragment (Probes) and pre-marking by H3K27me^3 in ES cells. Available upon request

References

1. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999;96:8681-8686.
2. Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, Whitehall VL, Leggett BA, Laird PW. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PLoS One 2009;4:e8357.
3. Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, Barker MA, Arnold S, McGivern A, Matsubara N, Tanaka N, Higuchi T, Young J, Jass JR, Leggett BA. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137-1144.
4. Minoo P, Moyer MP, Jass JR. Role of BRAF-V600E in the serrated pathway of colorectal tumourigenesis. J Pathol 2007;212:124-133.
5. Nagasaka T, Koi M, Kloor M, Gebert J, Vilkin A, Nishida N, Shin SK, Sasamoto H, Tanaka N, Matsubara N, Boland CR, Goel A. Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950-60, 1960.
6. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A 2007;%20;104:18654-18659.
7. Velho S, Moutinho C, Cirnes L, Albuquerque C, Hamelin R, Schmitt F, Carneiro F, Oliveira C, Seruca R. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis? BMC Cancer 2008;8:255.
8. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-793.
9. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology 2010;138:2073-2087.
10. Niv Y. Microsatellite instability and MLH1 promoter hypermethylation in colorectal cancer. World J Gastroenterol 2007;13:1767-1769.
11. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009;41:178-186.

12. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007;39:237-242.
13. Rada-Iglesias A, Enroth S, Andersson R, Wanders A, Pahlman L, Komorowski J, Wadelius C. Histone H3 lysine 27 trimethylation in adult differentiated colon associated to cancer DNA hypermethylation. Epigenetics 2009;4:107-113.
14. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 2007;39:232-236.
15. Kondo Y. Epigenetic cross-talk between DNA methylation and histone modifications in human cancers. Yonsei Med J 2009;50:455-463.
16. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. Epigenetic stem cell signature in cancer. Nat Genet 2007;39:157-158.
17. Xu X, Quiros RM, Gattuso P, Ain KB, Prinz RA. High prevalence of BRAF gene mutation in papillary thyroid carcinomas and thyroid tumor cell lines. Cancer Res 2003;63:4561-4567.
18. Yan PS, Chen CM, Shi H, Rahmatpanah F, Wei SH, Caldwell CW, Huang TH. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. Cancer Res 2001;61:8375-8380.
19. Wettenhall JM, Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics 2004;20:3705-3706.
20. Core Team RD. R: A Language and Environment for Statistical Computing. 2008.
21. Margaritis T, Lijnzaad P, van LD, Bouwmeester D, Kemmeren P, van Hooff SR, Holstege FC. Adaptable gene-specific dye bias correction for two-channel DNA microarrays. Mol Syst Biol 2009;5:266.
22. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological 1995;57:289-300.
23. Deng G, Chen A, Hong J, Chae HS, Kim YS. Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. Cancer Res 1999;59:2029-2033.
24. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997;25:2532-2534.
25. Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, Micklem G, Piano F, Snyder M, Stein L, White KP, Waterston RH. Unlocking the secrets of the genome. Nature 2009;459:927-930.
26. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006;125:301-313.
27. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, Orlov YL, Sung WK, Shahab A, Kuznetsov VA, Bourque G, Oh S, Ruan Y, Ng HH, Wei CL. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 2007;1:286-298.
28. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003;13:2129-2141.

29. Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 2009;10:697-708.
30. Lin J, Lai M, Huang Q, Ma Y, Cui J, Ruan W. Methylation patterns of IGFBP7 in colon cancer cell lines are associated with levels of gene expression. J Pathol 2007;212:83-90.
31. Carmona FJ, Esteller M. Epigenomics of human colon cancer. Mutat Res 2010;693:53-60.
32. Abel EV, Aplin AE. FOXD3 is a mutant B-RAF-regulated inhibitor of G(1)-S progression in melanoma cells. Cancer Res 2010;70:2891-2900.
33. de Keizer PL, Packer LM, Szypowska AA, Riedl-Polderman PE, van den Broek NJ, de BA, Dansen TB, Marais R, Brenkman AB, Burgering BM. Activation of forkhead box O transcription factors by oncogenic BRAF promotes p21cip1-dependent senescence. Cancer Res 2010;70:8526-8536.
34. Muggerud AA, Ronneberg JA, Warnberg F, Botling J, Busato F, Jovanovic J, Solvang H, Bukholm I, Borresen-Dale AL, Kristensen VN, Sorlie T, Tost J. Frequent aberrant DNA methylation of ABCB1, FOXC1, PPP2R2B and PTEN in ductal carcinoma in situ and early invasive breast cancer. Breast Cancer Res 2010;12:R3.
35. Lo PK, Lee JS, Liang X, Han L, Mori T, Fackler MJ, Sadik H, Argani P, Pandita TK, Sukumar S. Epigenetic inactivation of the potential tumor suppressor gene FOXF1 in breast cancer. Cancer Res 2010;70:6047-6058.
36. Nakshatri H, Badve S. FOXA1 in breast cancer. Expert Rev Mol Med 2009;11:e8.

Specific promoter methylation identifies different subgroups of *MLL*-rearranged infant Acute Lymphoblastic Leukemia, influences clinical outcome and provides therapeutic options

Blood (2009) 114(27):5490-8

Dominique J.P.M. Stumpel[1], Eddy H.J. van Roon[2§], Pauline Schneider[1§], Judith M. Boer[2], Paola de Lorenzo[3], Maria G. Valsecchi[3], Renee X. de Menezes[1,2], Rob Pieters[1] and Ronald W. Stam[1*]

[1]Department of Pediatric Oncology/Hematology, Erasmus Medical Center - Sophia Children's Hospital, Rotterdam, the Netherlands
[2]Leiden Genome Technology Center and Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands
[3]Department of Clinical Medicine, Prevention and Biotechnologies, Section of Medical Statistics, University of Milano-Bicocca, Monza, Italy
§ These authors contributed equally to this work.

Chapter 5

SPECIFIC PROMOTER METHYLATION IDENTIFIES DIFFERENT SUBGROUPS OF *MLL*-REARRANGED INFANT ACUTE LYMPHOBLASTIC LEUKEMIA, INFLUENCES CLINICAL OUTCOME AND PROVIDES THERAPEUTIC OPTIONS

Abstract

MLL-rearranged infant Acute Lymphoblastic Leukemia (ALL) remains the most aggressive type of childhood leukemia, displaying a unique gene-expression profile. Here we hypothesized that this characteristic gene-expression signature may have been established by potentially reversible epigenetic modifications. To test this hypothesis, we used Differential Methylation Hybridization (DMH) to explore the DNA methylation patterns underlying *MLL*-rearranged ALL in infants. The obtained results were correlated with gene-expression data to confirm gene silencing as a result of promoter hypermethylation. Distinct promoter CpG island methylation patterns separated different genetic subtypes of *MLL*-rearranged ALL in infants. *MLL* translocations t(4;11) and t(11;19) characterized extensively hypermethylated leukemias, whereas t(9;11)-positive infant ALL and infant ALL carrying wild-type *MLL* genes epigenetically resembled normal bone marrow. Furthermore, the degree of promoter hypermethylation among infant ALL patients carrying t(4;11) or t(11;19) appeared to influence relapse-free survival, with patients displaying accentuated methylation being at high relapse risk. Finally, we show that the demethylating agent zebularine reverses aberrant DNA methylation, and effectively induces apoptosis in *MLL*-rearranged ALL cells. Collectively these data suggest that aberrant DNA methylation occurs in the majority of *MLL*-rearranged infant ALL cases and guides clinical outcome. Therefore inhibition of aberrant DNA methylation may be an important novel therapeutic strategy for *MLL*-rearranged ALL in infants.

Introduction

While long-term survival rates in childhood Acute Lymphoblastic Leukemia (ALL) exceed 80%[1], the survival chances of infants (<1 year of age) still range between 20-50%[2]. Approximately 80% of infants with ALL carry chromosomal translocations involving the *MLL* gene[3], fusing the N-terminal portion of the *MLL* gene to the C-terminal region of one of its translocation partner genes. The most frequent *MLL* translocations among infant ALL patients are t(4;11), t(11;19) and t(9;11)[2,4], giving rise to the fusion proteins MLL-AF4, MLL-ENL and MLL-AF9. These chimeric MLL fusion proteins exhibit pronounced transforming capacities[5], and independently contribute to an unfavorable prognosis.[2, 6]

As a member of the trithorax gene family, *MLL* is involved in transcriptional regulation[7]. Therefore, structural alterations of this gene may be expected to affect its function, presumably leading to transcriptional deregulation. Not surprisingly, recent gene expression profiling studies characterized *MLL*-rearranged ALL as a unique type of leukemia that is genetically clearly separable from other ALL subtypes[8,9]. As epigenetic modifications affect gene expression patterns[10], we hypothesized that the specific gene expression profiles associated with *MLL*-rearranged infant ALL may well be driven by

epigenetic changes, which recently have been established to play important roles in the development and progression of leukemia[11]. The most widely studied epigenetic event in hematological malignancies constitutes transcriptional gene silencing by promoter CpG island hypermethylation[11,12]. This phenomenon either leads directly to the silencing of tumor suppressor genes, or indirectly to up-regulation of other genes, when silencing of certain regulatory genes relaxes the suppression on their target genes. Hence, genome-wide promoter hypermethylation potentially results in abnormal gene expression profiles that favor malignant transformation. For example, we recently demonstrated that *FHIT*, a putative tumor suppressor gene, is characteristically silenced in *MLL*-rearranged infant ALL cells by CpG hypermethylation, and that re-expression of this gene induced apoptosis in these cells[13].

Here we applied Differential Methylation Hybridization (DMH), an array-based technique that allows genome-wide screening of DNA methylation, using two different microarray platforms to explore the DNA methylation patterns underlying *MLL*-rearranged infant ALL. We show that different types of *MLL* translocations are associated with distinct patterns of DNA methylation, and we found that the degree of DNA methylation influences clinical outcome, identifying subgroups of *MLL*-rearranged infant ALL patients that may particularly benefit from therapeutic strategies containing demethylating agents.

Material and Methods

Patient samples

We studied 57 newly diagnosed infant ALL patients, enrolled in the international INTERFANT-99 treatment protocol[2] (patient characteristics listed in Supplemental table 1). Forty-four patients (77%) carried *MLL* translocations, and thirteen (23%) harbored untranslocated (wild-type) *MLL* genes. Among the *MLL* translocated patients, twenty one were positive for t(4;11), seventeen for t(11;19) and six patients carried translocation t(9;11). Written informed consent and institutional review-board approval were obtained for all patients. Whole normal bone marrow samples obtained from eight non-leukemic pediatric patients were included as controls. Leukemic cell isolation and enrichment to achieve more than 90% leukemic blasts, as well as DNA and RNA extractions were performed as described before[14].

Leukemia cell lines

RS4;11, SEMK2 and BEL-1 represent t(4;11)-positive precursor B-cell ALL cell lines. SEMK2 was originally derived from a 5-year-old girl at relapse[15] and was kindly provided by Dr. Scott Armstrong (Dana Farber Cancer Institute, Boston, Massachusetts, USA). BEL-1 was a generous gift from Dr. Ruoping Tang (University Laboratory, Paris, France).[16] RS4;11 was established from the bone marrow of a 32-year-old woman[17], and was, like all other cell lines used in this study, purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). KOPN-8 harbors translocation t(11;19) and was derived from a 3-month-old infant girl with B-cell precursor ALL. REH and TOM-1 represent precursor B-lineage ALL cells exhibiting a TEL-AML1 fusion, and a Philadelphia chromosome, respectively. JURKAT and HSB2 both are T-lineage ALL cell lines, and Kasumi-1 and MV4;11 are AML cell lines. Kasumi-1 carries the t(8;21) AML1-ETO fusion gene, and MV4;11 harbours *MLL* translocation t(4;11). All cell lines were maintained as suspension cultures in RPMI 1640 with L-Alanyl-L-Glutamine (Invitrogen) supplemented

with 10% FCS (Integro), 100 IU/ml penicillin, 100 µg/ml streptomycin, and 0.125 µg/ml fungizone (Invitrogen) at 37°C in humidified air containing 5% CO_2.

Differential Methylation Hybridization using CpG island microarrays

Differential Methylation Hybridization (DMH) was performed essentially as described by Yan et al. (Supplemental Methods)[18,19]. DMH was applied on two different CpG island microarray platforms with limited overlap in CpG island probes. The first was the custom spotted 9K microarray chip developed by Huang and co-workers, containing 8,640 *Mse*I fragment probes[18]. In addition, we also used the first commercially available genome-wide CpG island microarrays (Agilent Technologies, Santa Clara, USA). These high-resolution microarrays contain 243,497 60-mer oligonucleotide probes, including 67,487 CpG island probes located in or near gene promoters. For the present study, only these probes located in gene promoters were used. Due to restricted availability of patient material, DNA methylation profiling using the Agilent microarrays was performed in 49 of the 57 infant ALL patients.

Gene expression profiling using Affymetrix GeneChips

Gene expression profiles were generated for t(4;11)-positive (n=15) and t(11;19)-positive (n=14) infant ALL cases, using the same samples for which DNA methylation profiles were already produced on Agilent microarray chips. Expression profiles were also generated for whole healthy pediatric bone marrow samples, however, these did not correspond to the samples in which the DNA methylation patterns were determined. RNA processing, microarray hybridization (HU133 plus 2.0 Affymetrix GeneChips), and washing steps were performed according to the manufacturer's protocol (Affymetrix, Santa Clara, CA, USA). The infant ALL gene expression data and DNA methylation data presented in this study have been deposited in the NCBI Gene Expression Omnibus and are accessible via the GEO Series accession number GSE18400.

In vitro cytotoxicity assay and exposure to zebularine

In vitro sensitivity of leukemia cell lines to the demethylating agent zebularine[20,21] was determined by four-day MTT-assays as described previously[22]. Zebularine was a generous gift from Dr. Victor E. Marquez (National Cancer Institute of Frederick, Frederick, Maryland, USA). To study the effects of demethylation on *MLL*-rearranged ALL cells, the cell lines SEMK2 and RS4;11 were cultured for 10 days in the presence or absence of 100 µM zebularine.

Statistical analyses

Normalization of the CpG island microarray data was performed using global locally weighted scatterplot smoothing (loess) normalization[23], and differentially methylated CpG islands were identified using the linear models for microarray data (limma) package in the R statistical environment (R Development Core Team, 2007) (Supplemental Methods) [24]. The resulting list of p-values was corrected for multiple testing by the false discovery rate (FDR) step-up procedure of Benjamini & Hochberg[25]. An FDR-adjusted p-value <0.01 was regarded significant. As a measure of internal validation for the subtype-specific methylation signatures, permutation testing (global test)[26] was applied to evaluate whether genes were significantly associated with a certain type of *MLL* translocation. For this, the tendency of repeated re-assignment of individual samples to their original cluster

was assessed. (Supplemental Methods)

Relapse-free survival was computed with the Kaplan Meier estimator. The duration of relapse-free survival was defined as the time from diagnosis until the date of leukemia relapse or the last follow-up. The probability of relapse in complete remission was estimated by applying the cumulative incidence estimator. The log-rank test was used to compare outcomes between different patient groups, and a one-step Cox model was applied to estimate the hazard of relapse for these patients, adjusting for already established risk stratification according to the international INTERFANT-06 treatment protocol. (Supplemental Methods).

Results

Unsupervised analysis based on DNA methylation patterns separates different infant ALL subtypes

Using Differential Methylation Hybridization (DMH) on two different microarray platforms, genome-wide promoter DNA methylation profiles were generated for infant ALL patients carrying *MLL* translocations t(4;11), t(11;19) or t(9;11), and infant ALL patients bearing wild-type *MLL* genes. To explore whether these samples showed leukemia-specific increases in promoter CpG island methylation, these profiles were compared with DNA methylation patterns obtained from bone marrow samples derived from healthy children. Initially, we performed a principal component analysis (PCA), using all CpG island probes present on each array without any selection. Based on the first three components of the PCA, which explain 41.8% (9K chip) and 32.2% (244K chip) of the total variance, the patient samples were visualized (Figure 1). Interestingly, for both microarray platforms, this unsupervised analysis separated two major groups. Infant ALL samples that carry t(9;11) or wild-type *MLL* genes clustered together with normal bone marrow samples, whereas infant ALL samples carrying t(4;11) or t(11;19) clustered tightly together separately from the other samples. Although the cluster comprising t(9;11)-positive, untranslocated infant ALL samples and normal bone marrow samples appeared more heterogeneous, it has to be taken into account that this cluster consists of three different types of samples. Moreover epigenetic heterogeneity is already present among the normal bone marrow samples. Finally, we emphasize that this analysis is completely unguided.

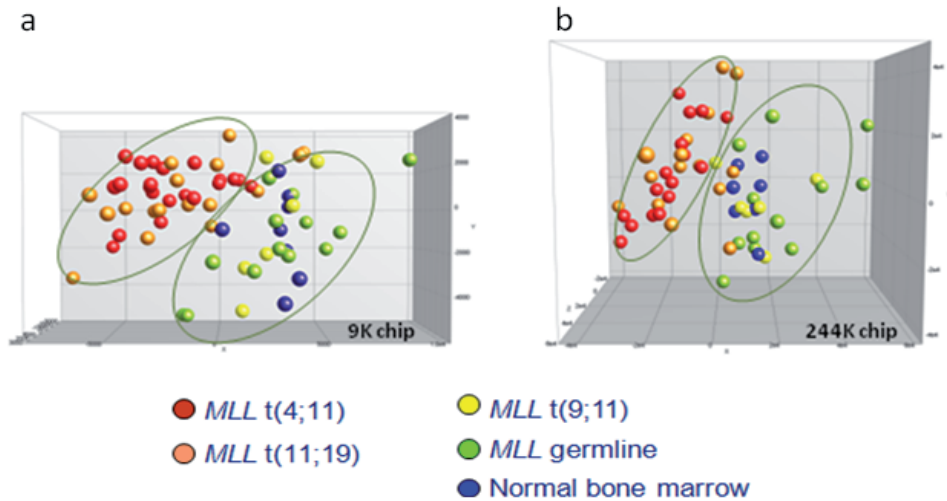


Figure 1 - Unsupervised clustering analysis of DNA methylation in infant ALL. Principal Component Analyses (PCA) of the CpG island methylation data from infant ALL patients and normal bone marrows using all probes present on each microarray platform. Each case is color-coded indicating the specific infant ALL subgroups. **a-1.** shows data from the custom spotted 9K CpG island microarray. t(4;11) (n=21; red), t(11;19) (n=17; orange), t(9;11) (n=6; yellow), *MLL* wild-type ALL (n=13, green) and normal bone marrow (n=8; blue). **a-2.** shows data from the commercially available 244K CpG island microarray (Agilent). t(4;11) (n=16; red), t(11;19) (n=15; orange), t(9;11) (n=6; yellow), *MLL* wild-type ALL (n=12; green) and normal bone marrow (n=7; blue). Due to restricted availability of patient material, Agilent DNA methylation profiles were generated for 49 infant ALL patients and 7 normal bone marrow samples.

Specific DNA methylation patterns further separate the different infant ALL subtypes

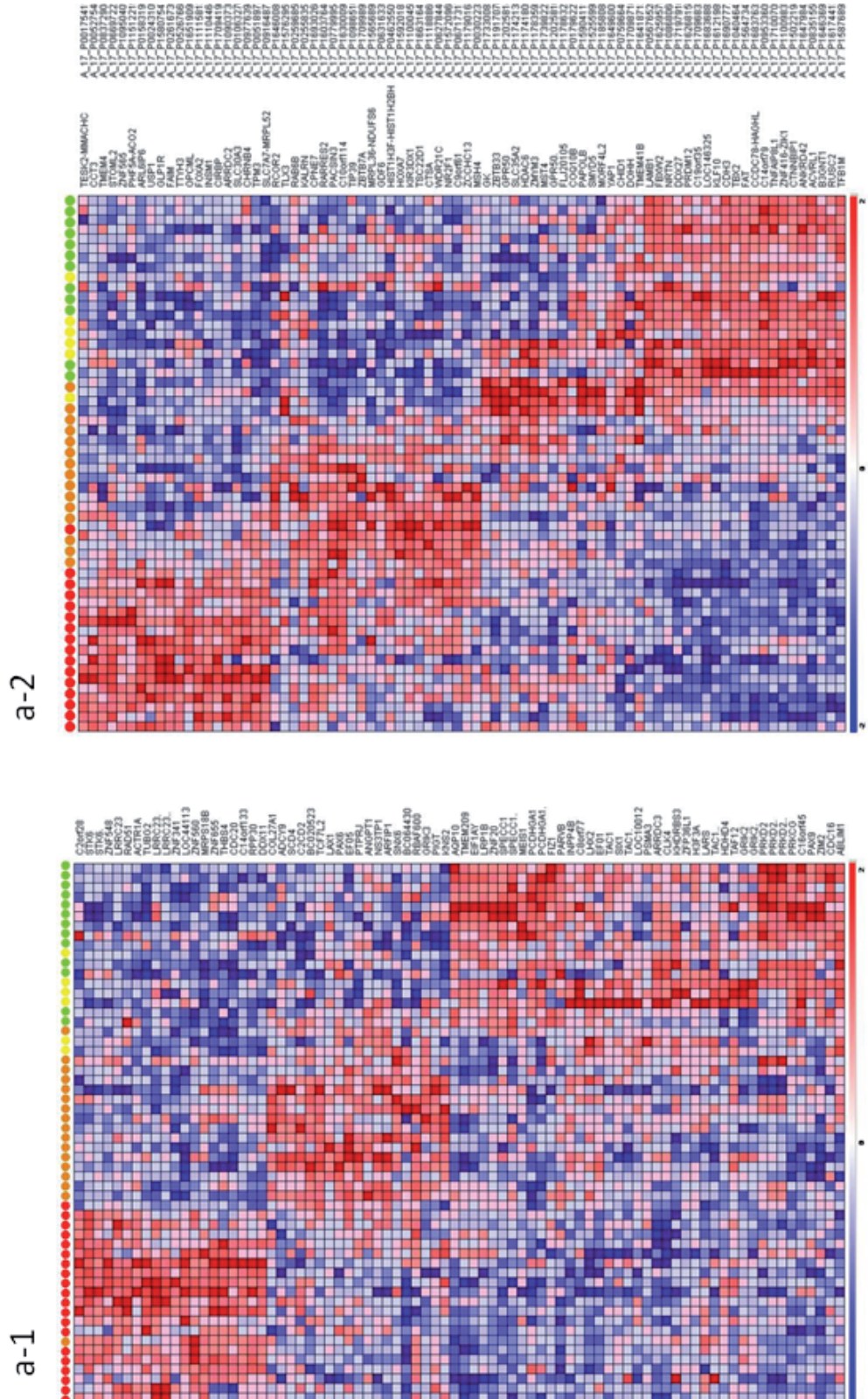
Subsequently, to explore whether specific DNA methylation profiles could define the genetic subgroups of infant ALL more accurately, the 20 most discriminative hypermethylated genes for each group (as compared with all other relevant subgroups combined) were selected. For both microarray platforms the gene names, log-fold changes, and p-values are listed in the Supplemental data (tables 2S and 3S). Permutation testing validated the robustness of the subtype-specific methylation signatures. Using the selected genes, we generated heatmaps in which both the genes and samples were clustered hierarchically (Euclidean distance and complete linkage) (Figure 2a). This semi-supervised analysis revealed that *MLL* t(4;11) and t(11;19)-positive patients could clearly be separated from one another and from the other samples. In contrast, hypermethylated genes that unambiguously separate t(9;11)-positive samples from *MLL* wild-type (untranslocated) infant ALL samples, could not be identified. Moreover, the most significantly hypermethylated genes shared by t(9;11)-positive and wild-type *MLL* samples were also methylated in healthy bone marrow samples (Supplemental figure 1S), and therefore likely reflect normal methylation in healthy hematopoietic cells. Importantly, these genes are hypomethylated in infant ALL harbouring translocation t(4;11) or t(11;19).

Next, PCA was used to better visualize these different clusters, emphasizing the separation of the samples into the three expected groups characterized by t(4;11), t(11;19) or t(9;11) together with translocation-negative infant ALL (Figure 2b). When included, the normal

bone marrow samples remained within the cluster comprising samples carrying t(9;11) or wild-type *MLL* genes (Figure 2c). In concordance with this, no significant aberrant DNA methylation could be detected in t(9;11)-positive or untranslocated (wild-type *MLL*) infant ALL, when separately compared with normal bone marrow.

Correlation between promoter methylation and gene expression

Given the aberrant methylation patterns in t(4;11)- and t(11;19)-positive infant ALL samples, we investigated the effects of promoter hypermethylation (Agilent platform) on gene expression (Affymetrix platform) of corresponding genes. Compared with normal bone marrow samples, infant ALL cells carrying t(4;11) displayed a total of 794 hypermethylated CpG island probes (FDR<0.01), and 75 probes were significantly hypermethylated in t(11;19)-positive infant ALL (FDR<0.01). From these analyses, the most significantly hypermethylated genes were selected. Gene names, log-fold changes in methylation, and p-values for these genes are listed in the Supplemental data (tables 4S and 5S). Next, DNA methylation array data was compared with gene expression profiles from the same samples, and visualized as heatmaps and PCA plots (Figure 3). Promoter hypermethylation and down-regulated gene expression correlated for ~90-95% of the genes in both t(4;11)-positive and t(11;19)-positive infant ALL. However, for the remaining genes we observed the opposite; despite extensive hypermethylation, these genes were higher expressed in leukemic samples than in normal bone marrow.


a-1
a-2

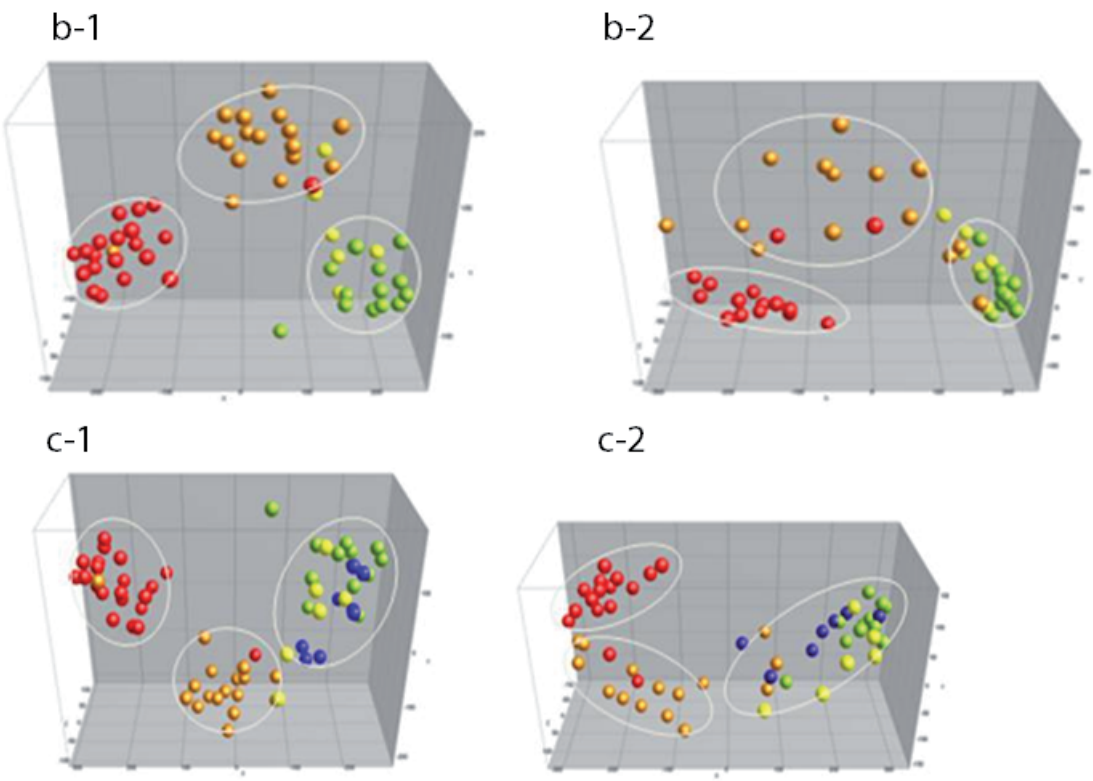


Figure 2 - Infant ALL subtype-specific CpG island hypermethylation. **a1-2.** Heatmaps showing the 20 most significantly hypermethylated probes for each infant ALL subtype. Columns represent patient samples and rows represent genes. Relative DNA methylation levels are shown in red (high) and blue (low). Genes and samples were ordered using hierarchical cluster analysis (Euclidean distance, complete linkage), and gene identifiers are listed at the right. **b1-2.** Principal component analyses (PCA) separating t(4;11) (red), t(11;19) (orange), t(9;11) (yellow) and *MLL* wild-type infant ALL (green). **c1-2.** shows the PCA when normal pediatric bone marrow samples (blue) are included in the analysis. **2-1.** shows data from the custom 9K CpG island microarray. **2-2.** shows data from the commercially available 244K CpG island microarray (Agilent).

The degree of methylation influences clinical outcome in MLL-rearranged infant ALL

In both the heatmaps and PCA plots that represent the most significantly hypermethylated genes among t(4;11)-positive and t(11;19)-positive infant ALL, a clustering of patient samples into two subgroups appeared. Ostensibly, one cluster represents patient samples that, at least for the selected genes, seem to be more densely hypermethylated than the samples in the other cluster. (Figure 3) To better visualize this difference in degree of methylation we plotted the normalized methylation log-ratios of the genes. This semi-quantitative representation of the data indeed confirmed differences in the degree of methylation between both clusters (Figure 2S). To explore the clinical relevance of these subgroups, we computed risk of relapse statistics for these patient groups (Figure 3S). Four patients received bone marrow transplantation (BMT) in complete remission. For these patients data were censored at BMT. One patient died before the start of treatment (referred to as early death), and was excluded from further analyses. Twelve out of 16 (75%) patients from the "heavily" methylated subgroup had a relapse after achieving complete remission, whereas among the "lightly" methylated patients, relapses occurred in 5/12 (42%) of the cases. The cumulative incidence of relapse at 1 year after diagnosis was significantly ($p<0.05$) different for the "heavily" and "lightly" methylated subgroups with incidences of 52.5 (SE: 13.7) and 35.7 (SE: 15.5) respectively. The number of patients that could be included in these analyses is not sufficient to evaluate the impact of the degree of DNA methylation adjusted for known prognostic factors (like age, white blood cell count and the *in vivo* response to prednisone) separately. Therefore we used the INTERFANT-06 risk stratification which represents a combination of these factors (see: Supplemental Methods)[2]. Although these results must be interpreted with caution, the Cox regression model indicates that heavy methylation confers an increased risk of relapse (hazard ratio 5.77, 95% CI 1.57-21.2, $p=0.01$) (table 6S). The separate clustering of these two patient groups, however, did not appear in the gene expression profiles (Figure 3). This implies that the grouping of these patients and the observed variance in relapse-free survival rather reflects progressive accumulation of genome-wide methylation, than direct differences in gene expression. In line with this hypothesis we show that the division into "heavy" and "light" methylation remains present when all significantly hypermethylated probes are used in a semi-quantitative representation (Figure 2S).

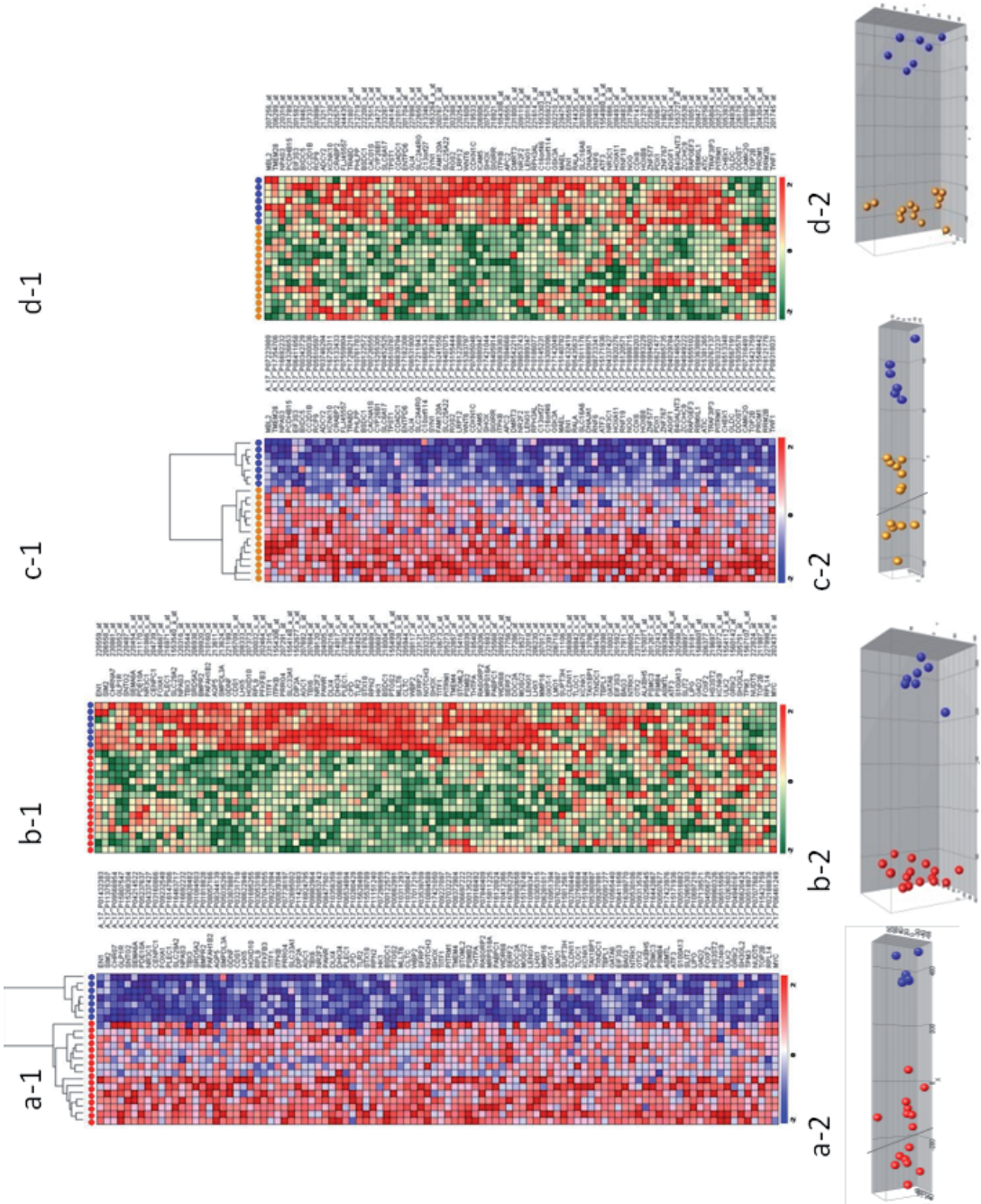


Figure 3 - Correlation between CpG island methylation and gene expression. **a.** Heatmap and PCA showing the most significantly hypermethylated genes in t(4;11)-positive infant ALL (red dots) compared with normal bone marrows (blue dots). Relative DNA methylation levels are shown in red (high) and blue (low). **b.** Heatmap and PCA showing the corresponding gene expression levels from the same genes and samples as presented in Figure 3A. Relative gene expression values are shown in red (high) and green (low). Similarly, the **c.** CpG methylation data and **d.** gene expression data are presented for t(11;19)-positive infant ALL samples (orange dots) as compared with normal bone marrows (blue dots).

MLL t(4;11)-*positive cell lines as models for demethylation*

DNA methylation patterns of two t(4;11)-positive precursor B-cell ALL cell lines (i.e. RS4;11 and SEMK2) were compared with the profiles from the t(4;11)-positive infant ALL samples. Nearly 50% of the 100 most significantly hypermethylated genes in t(4;11)-positive infant ALL were also hypermethylated in these cell lines (Figure 4a). Representing reasonable models for t(4;11)-positive infant ALL samples, we next studied the effects of demethylation on these genes by comparing DNA methylation profiles of these cell lines before and after a 10-day exposure to 100 μM of the demethylating agent zebularine. In the SEMK2 and RS4;11 cell lines, respectively 72% (33/46) and 59% (27/46) of the hypermethylated genes showed notable decreases in methylation upon exposure to zebularine. The genes display varying degrees of drug-induced demethylation (Figures 4b-c). For some of the genes the methylation status could be restored to nearly normal levels as observed in healthy hematopoietic cells.

Specific zebularine sensitivity in MLL-rearranged ALL cells

To further investigate the sensitivity of ALL cells to *in vitro* demethylation, cytotoxicity assays were performed using escalating dosages of zebularine. Also two AML cell lines were added to the data set. AML cells (with or without an *MLL* translocation) seem to be less sensitive to the demethylating agent zebularine than *MLL*-rearranged ALL cells, but the *MLL*-rearranged AML cell line MV4-11 does appear more sensitive than the t(8;21)-positive AML cell line Kasumi-1. Clearly, *MLL*-rearranged ALL cells were significantly more sensitive to zebularine than the other cell lines ($p<0.01$) (Figures 5A-B). As shown in figure 5B, on average the IC_{50} value (i.e. the concentration inhibitory to 50% of the cells) in *MLL*-rearranged ALL cells was ~50 μM, whereas zebularine failed to reach an IC_{50} value in other types of ALL cell lines.

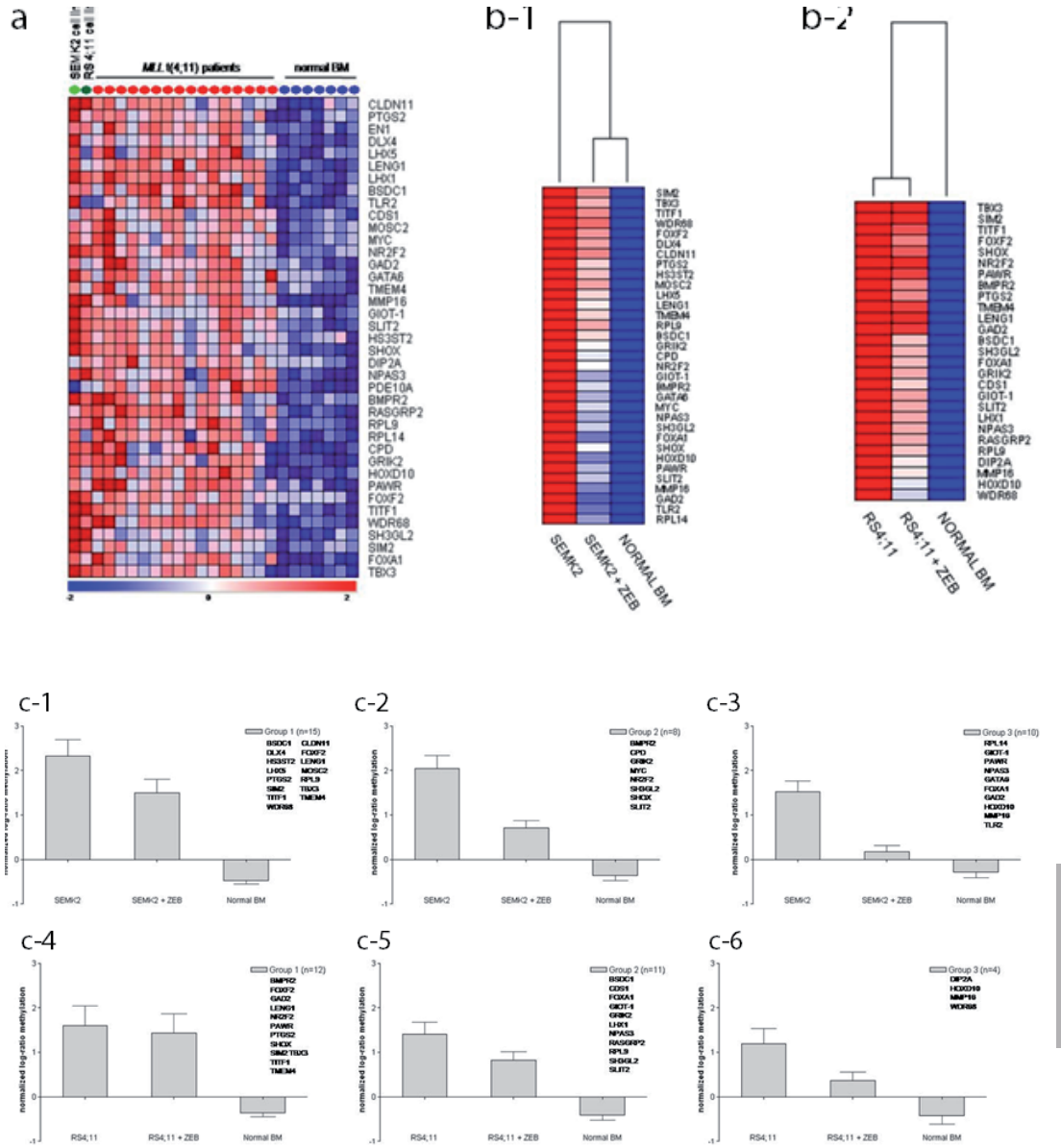


Figure 4 - ALL cell lines as models for (de)methylation **a.** Heatmap showing methylation levels in the t(4;11)-positive B-ALL cell lines SEMK2 (light green dot) and RS4;11 (dark green dot) of genes most significantly methylated in t(4;11)-positive infant ALL patients (red dots) as compared with normal bone marrows (blue dots). **b.** Heatmaps showing methylation levels of these genes after exposure to zebularine. These methylation levels were compared with the average methylation levels as determined from normal bone marrow samples (n=7). **b1** SEMK2 cell line and **b2.** RS4;11 cell line **c1-6.** Graphs displaying the mean and the standard error of the mean (SEM) of changes in methylation levels after zebularine exposure. Genes were divided into three groups for each cell line according to the degree of responsiveness to zebularine. **c1-3.** SEMK2 cell line and **c4-6.** RS4;11 cell line.

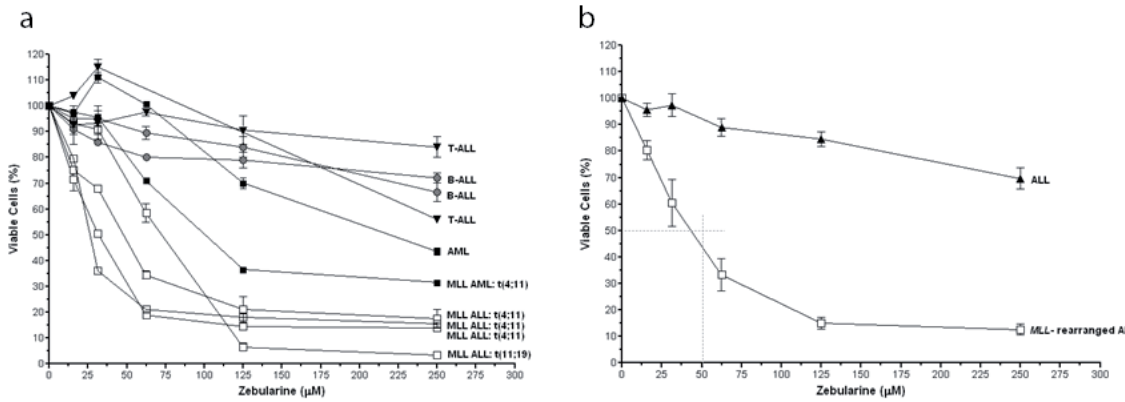


Figure 5 - *In vitro* cytotoxicity to zebularine. **a.** Dose-response curves showing the *in vitro* cytotoxic response to zebularine in individual leukemia cell lines with or without *MLL* rearrangements, or **b.** the mean cytotoxic response for *MLL*-rearranged ALL cell lines (n=4) and for the other ALL cell lines (n=4). Error bars represent standard error of the mean (SEM). The differences between the means of the groups were statistically analyzed using the 2-tailed Student *t* test ($p<0.01$ for each concentration used).

Discussion

We here present the first global view of the DNA methylome in infant *MLL*-rearranged Acute Lymphoblastic Leukemia (ALL). *MLL*-rearranged infant ALL represents an aggressive and difficult to treat type of leukemia characterized by a unique gene expression profile, that clearly separates this malignancy from other ALL subtypes[8, 9]. Since epigenetic modifications directly influence gene expression patterns[10], we hypothesized that specific DNA methylation patterns may underlie the characteristic gene signature as observed for *MLL*-rearranged infant ALL. Our data largely support this hypothesis, as the majority of *MLL*-rearranged infant ALL cases (i.e. those characterized by t(4;11) or t(11;19)) represent hypermethylated leukemias, whereas t(9;11)-positive and *MLL* translocation-negative (wild-type *MLL*) infant ALL display DNA methylation patterns that closely resemble that of normal bone marrow. Moreover, distinct leukemia-specific DNA methylation patterns could be identified for the different *MLL*-rearranged infant ALL subtypes as defined by the type of *MLL* translocation or absence of such translocations. Interference of non-leukemia related epigenetic differences in DNA methylation (such as age, sex-specific differences in methylation and differences related to B-cell maturation stages of leukemic cells) with our results could be excluded (Supplemental Results, tables 7S and 8S).

Thus, the presence as well as the patterns of aberrant DNA methylation in infant ALL appear, at least to some extent, dependent on the presence and type of *MLL* fusion, which may reflect a mechanism proposed recently[27]. Apart from DNA methylation, a second component of the epigenetic code involves histone modifications[28], shaping the chromatin in an open (transcriptionally active) or closed (inactive) conformation. An inactive chromatin state usually is associated with hypermethylated CpG promoter regions, whereas active chromatin marks, such as H3K4 trimethylation and H3K79 dimethylation, denote unmethylated promoters, allowing transcription. Interestingly, the *MLL* gene itself has specific histone methyltransferase activity[29,30], which is lost during fusion of the *MLL* gene to one of its translocation partners. Therefore, MLL fusions can be expected to result in altered chromatin structures due to aberrant histone modifications. Recently,

Krivtsov and Armstrong (2007) proposed that the recruitment of different histone methyltransferases by different MLL fusion proteins, may indeed result in inappropriate histone modifications directed by the *MLL* fusion partner[27]. Given the sound interplay between histone modification and CpG island methylation, this proposed influence of different *MLL* fusion genes on histone modifications, and the apparent influence of the *MLL* fusion partner on DNA methylation as shown in the present study, are presumably linked. In addition, the paper by Krivtsov et al.(2007)[27],as wells as the study by Mueller and co-workers (2007)[31] demonstrated the recruitment of a transcriptional elongation complex to *MLL* target genes, resulting in gene activation (i.e. expression). These studies suggest that MLL fusion proteins trigger or maintain the leukemia by the activation of specific target genes. In contrast, our present study shows that, apart from specific gene activation, *MLL*-rearranged ALL is also characterized by severe gene inactivation, which may well be driven by the same MLL fusion. *MLL*-rearranged ALL cells typically mirror highly immature B-cells. Possibly, the MLL fusion ignores the activation of many genes that should have been activated (by wild-type *MLL*) at this stage of B-cell development, and necessary for proper differentiation towards mature and functional B-cells. This would suggest that our observed patterns of gross genome-wide DNA methylation is in favor of blocking B-cell differentiation, while simultaneously the MLL fusion activates several (proto-onco)genes in favor of uncontrolled cell proliferation and survival. Alternatively (or additionally), inappropriate activation of certain genes by the MLL fusion, may in turn induce abnormal inactivation (silencing) of several other genes. However, these proposed mechanisms are highly speculative and remain to be confirmed.

Nonetheless, we can conclude from our data that MLL-AF4 and MLL-ENL represent MLL fusion proteins that both alter histone modifications that result in strongly altered DNA methylation patterns. The differences found in DNA methylation patterns between MLL-AF9 and MLL-ENL may then seem surprising given the apparent common mechanism of transformation involving the recruitment of *DOT1L* as put forward by others [27,31]. Surprisingly, the MLL-AF9 fusion did not lead to significant aberrant DNA methylation in infant ALL. This suggests that oncogenic transformation in t(4;11)- and t(11;19)-positive infant ALL patients may be facilitated or largely driven by gross epigenetic changes, whereas t(9;11)-positive infant ALL cells presumably transform via alternative mechanisms. In concordance with this is that t(9;11)-positive ALL patients characteristically seem to be different from other *MLL*-rearranged infant ALL patients. For example, t(9;11)-positive infant ALL is typically diagnosed at a later stage during infancy and usually is characterized by a more mature immunoglobulin gene rearrangement pattern (immunophenotype) than t(4;11)- and t(11;19)-positive infant ALL[2,4]. On the other hand, no significant differences in survival exist between infant ALL patients carrying either t(4;11) or t(11;19) and patients with t(9;11).[2]

Studying the genes most significantly hypermethylated in t(4;11)- and t(11;19)- positive infant ALL samples, we found that the expression of the vast majority of these genes (~90-95%) was indeed down-regulated. Among the hypermethylated genes we found genes that were previously described to be silenced due to DNA hypermethylation in *MLL*-rearranged ALL, such as the tumor suppressor gene *FHIT*[13] and the *DLX3* gene[32], demonstrating the integrity of our data. Moreover, most of these genes responded well to exposure to the demethylating agent zebularine in t(4;11)-positive cell line models. Among the most significantly hypermethylated genes for either t(4;11)-positive or t(11;19)-

positive infant ALL, a limited overlap was observed. Nevertheless, global gene ontology analysis showed that most of the down-regulated genes in both subgroups are involved in transcriptional regulation (table 9S). This pronounced epigenetic deregulation of the transcriptional machinery may indeed have contributed to the unique gene expression profile characteristic for *MLL*-rearranged ALL[8,9]. Yet, this would not be true for t(9;11)-positive infant ALL, as no aberrant promoter hypermethylation was observed in these samples. This apparent contradiction, however, is easily explained by the fact that most of the published *MLL*-specific gene expression signatures, including the signatures reported by Armstrong et al. (2002) are predominantly based on t(4;11)- and t(11;19)-positive samples[8]. Therefore gene expression profiling studies including t(9;11)-positive infant ALL samples may well come to demonstrate that profiles associated with t(9;11) are different from those obtained in t(4;11)- and t(11;19)-positive samples.

Remarkably, about 5% of the most significantly hypermethylated genes in t(4;11)- and t(11;19)-positive infant ALL remained highly expressed. This observation controverts the dogma that promoter methylation per definition induces suppression of gene expression. However, Weber and co-workers recently nuanced this dogma by elegantly demonstrating the influence of promoter CpG density on the ability to induce transcriptional repression[33]. Therefore, these methylated but highly expressed genes may well exhibit promoters containing weak CpG islands (i.e. a low or intermediate CpG density), unable to repress transcription even when methylated. Another possible explanation for this would again be the involvement of the MLL fusion protein, which may have induced activating histone modifications on otherwise inactive regions in the chromatin associated with promoter methylation. In turn, this newly acquired open chromatin state may have overruled the relatively weak DNA methylation, allowing transcription despite earlier established epigenetic silencing. If so, this group of genes may well represent potential therapeutic targets directly influenced by the MLL fusion itself.

Most of the genes that were methylated in t(9;11)-positive infant ALL and infant ALL carrying wild-type *MLL* genes were also methylated in normal bone marrow. Presumably, these represent genes that were already silenced in normal hematopoietic cells, but became hypomethylated in t(4;11) and t(11;19) positive infant ALL cells. Interestingly, among these were several genes with oncogenic potential, such as *CDH3*, *TBX2*, *ERCC1* and *NPR2* (figures 2A-B), that have been reported to be involved in proliferation, tumor aggressiveness and prognosis in a wide range of human cancers[34,35]. Interestingly, among these hypomethylated genes also appeared the *HOXA9* gene which was previously described to be protected from methylation by the MLL-fusion itself[36]. Thus, the present study not only characterizes epigenetically down-regulated genes, but also identifies proto-oncogenes that may be inappropriately expressed in t(4;11)- and t(11;19)-positive *MLL*-rearranged ALL in infants. Obviously, such genes represent yet another set of candidate target genes for future therapeutic intervention.

Of main therapeutic interest is our finding that the degree of DNA methylation among t(4;11)- and t(11;19)-positive infant ALL patients is related to relapse-free survival, with patients presumably carrying heavily methylated genomes being at an increased risk of relapse. Therefore, these children in particular should be considered candidates for therapies including inhibitors of DNA methylation, especially since we here show that *MLL*-rearranged ALL cells are highly sensitive to zebularine *in vitro*. The authors believe that this increased sensitivity to demethylation is rather based on the presence of a general

methylator phenotype (i.e. globally deregulated DNA methylation) than on the actual re-expresssion of a fixed number of hypermethylated genes. Apparently, genome-wide demethylation is sufficient to cause *MLL*-rearranged ALL cells to undergo apoptosis. This is in concordance with the identification of a heavily and a lightly methylated subgroup of *MLL*-rearranged infant ALL, which is also based on a widespread phenotype with more or less pronounced levels of DNA methylation that are in fact not visible at the gene-expression level. In conclusion, the findings presented here urgently require gene per gene validation studies and mandate additional studies using demethylating agents in the currently only available genuine mouse model for *MLL*-rearranged ALL, recently described by Krivtsov et al.[37]

Acknowledgments
The authors wish to express gratitude to the members and participating hospitals of the INTERFANT-99 study for supporting our research by providing leukemic samples. Members of INTERFANT-99 are: Campbell, M. (PINDA), Felice, M. (Argentina), Ferster, A. (CLCG), Hann, I. and Vora, A. (UKCCSG), Hovi, L. (NOPHO), Janka-Schaub, G. (COALL), Li, CK. (Hong Kong), Mann, G. (BFM-A), LeBlanc, T. (FRALLE), Pieters, R. (DCOG), de Rossi, G. and Biondi, A. (AIEOP), Rubnitz J. (SJCRH), Schrappe, M. (BFM-G), Silverman, L. (DFCI), Stary, J. (CPH), Suppiah, R. (ANZCHOG), Szczepanski, T. (PPLLSG), Valsecchi, M. and de Lorenzo, P. (CORS). Furthermore, T.H. Huang and P. Yan are gratefully acknowledged for contributing the 9K CpG island clone library and protocols for differential methylation hybridization.

Supplementary data

Supplemental Results

Exclusion of non–leukemia-related epigenetic differences between subtypes of infant ALL

To test whether non–leukemia-related epigenetic differences were present among the sample population that may compromise our findings, we performed several comparisons deducting this possibility. Since leukemic samples were either obtained as peripheral blood or as bone marrow, we first compared DNA methylation patterns in bone marrow samples with samples derived from peripheral blood. No significant differences in promoter CpG island methylation were observed. From this, we concluded that samples of both sources could legitimately be used together in our analyses.

Next, and similarly, sex differences were assessed. A female predominance (62% vs. 38%) was identified among the *MLL*-rearranged infant ALL subgroups, whereas a male predominance (61% vs. 39%) was present in the translocation-negative (wild-type *MLL*) infant ALL group. 176 probes were differentially methylated between male and female patients (FDR<0.01) according to the Agilent microarray platform (Tables S7 and S8). The 121 probes corresponding to 75 genes, hypermethylated in females were all located on the X chromosome. Conversely, most of the 55 probes, corresponding to only 5 genes hypermethylated in male individuals were located on the Y chromosome. None of these genes appeared to interfere with the infant ALL subtype-specific methylation signatures. In general, infant ALL represents a highly immature type of precursor B-cell leukemia with t(4;11)- and t(11;19)-positive infant ALL most often

displaying the most immature phenotypes.[1] We therefore investigated whether the identified DNA methylation profiles to some extent reflected different maturation stages of early B-cell development. For this, we compared samples with a pro–B-cell ($CD34^+$, $CD19^+$, ($CD10^-$)) leukemia with the more mature pre–B-cell ($CD34^-$, $CD19^+$, $Cylg\mu^+$) leukemia within the group of samples carrying t(11;19), t(9;11) or wild-type *MLL* genes. Infant ALL samples carrying t(4;11) were excluded because they nearly all represented pro–B-cell leukemias and thus would strongly bias these analyses. No significant differences were found between pro-B and pre-B infant ALLs.

Finally, no significant differences in promoter CpG island methylation could be observed when DNA methylation patterns from infants diagnosed below 6 months of age were compared with those from infants diagnosed between 6 months and 1 year or age.

Taken together these comparisons indicate that the DNA methylation patterns as presented in this study represent leukemia-specific profiles that are unlikely to be influenced by differential CpG island methylation induced by non-leukemic factors.

Supplemental Methods

Differential methylation hybridization using CpG island microarrays

Differential methylation hybridization (DMH) was performed as described by Yan et al.[2] Briefly, 0.5 ug of high-quality genomic DNA was digested using the restriction enzyme *Mse*I (New England Biolabs). Next, unphosphorylated linkers were ligated to the digested fragments, which were then sequentially digested with two methylation-sensitive restriction enzymes (*Bst*UI and *Hpa*II, New England Biolabs). These second digestions eliminate unmethylated fragments, enriching the samples for methylated sequences. The digested linker-ligated DNA was then used as a template for polymerase chain reaction (PCR) amplification (20 cycles of 97°C for 1 min and 72°C for 3 min, final extension at 72°C for 10 min), generating methylated amplicons. Using the BioPrime Array-CGH Genomic Labeling kit (Invitrogen, Carlsbad, USA), amino-allyl dUTPs were incorporated into the amplicons, allowing the amplicons to be labeled with the fluorescent dyes Cy5 (patient samples) and Cy3 (common reference samples). The common reference for all samples was a commercially available genomic DNA pool derived from five healthy males and five healthy females (Promega Benelux BV, Leiden, the Netherlands). Hybridization and washing were performed according to DeRisi[3] for the custom 9K microarray chips, and for the 244K microarray chips (Agilent), the Agilent ChIP-on-chip protocol version 9.0 was used. Hybridized slides were scanned with a 2565 AA DNA microarray scanner (Agilent Technologies), and the acquired images were analyzed using the GenePix Pro 6.0 software or the Agilent Feature Extraction 9.5.3 software.

Statistical analyses and software

Analyses of differential methylation

Differentially methylated CpG islands were identified using linear models for microarray data (limma).[4] These models use an empirical Bayes approach to moderate the standard errors of the estimated standardized log-fold changes by borrowing information across genes. This results in more stable assumptions and enhanced power, especially when group sizes are small.[5]

Permutation testing

As a measure of internal validation for the subtype-specific methylation signatures, permutation testing (global test) [6] was applied to evaluate whether genes were significantly associated with a certain type of *MLL* translocation. Therefore the tendency of repeated re-assignment of individual samples to their original cluster was assessed. As a default 10.000 permutations were used to calculate p-values.

Gene expression

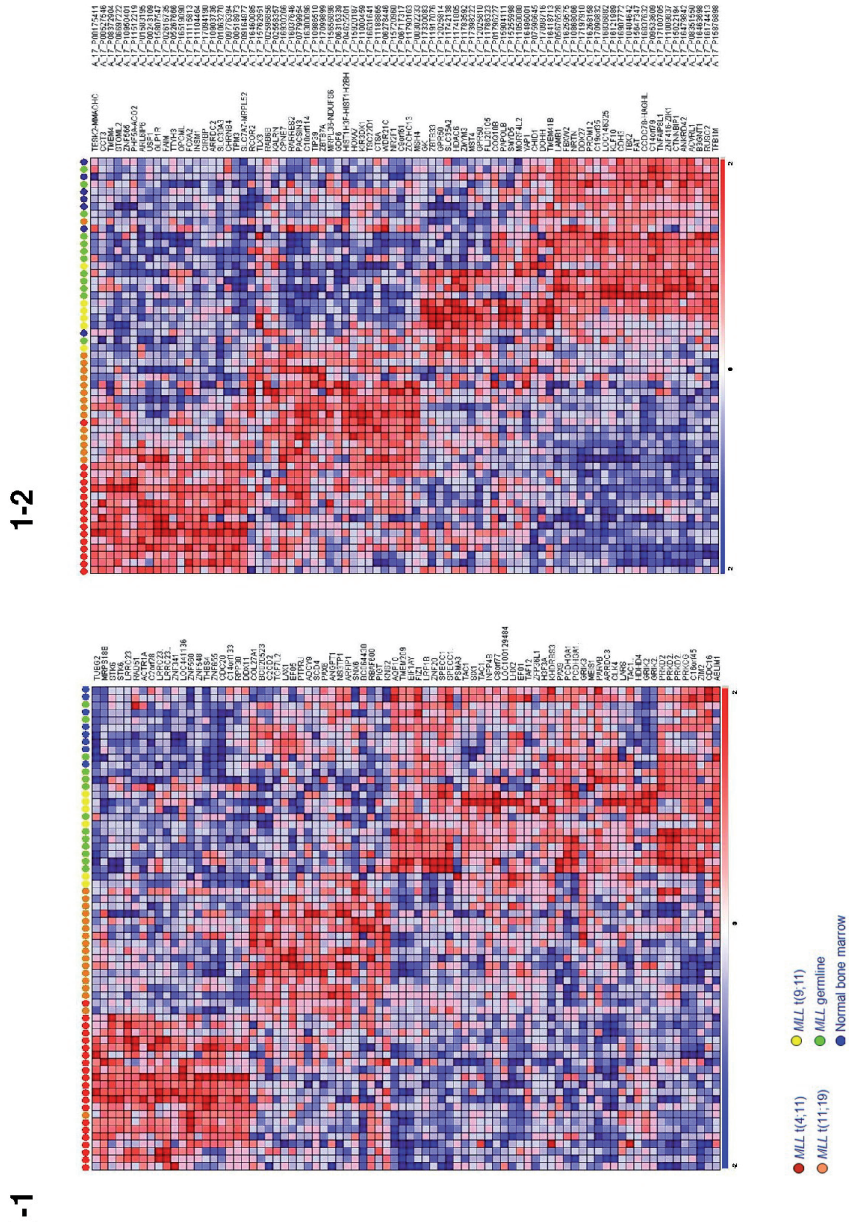
Gene expression values were calculated using Affymetrix Microarray Analysis Suite® (MAS) 5.0.2 software. Unscaled expression signals were normalized using variance stabilization and normalization (vsn).[7]

Risk stratification according to the INTERFANT-06 treatment protocol

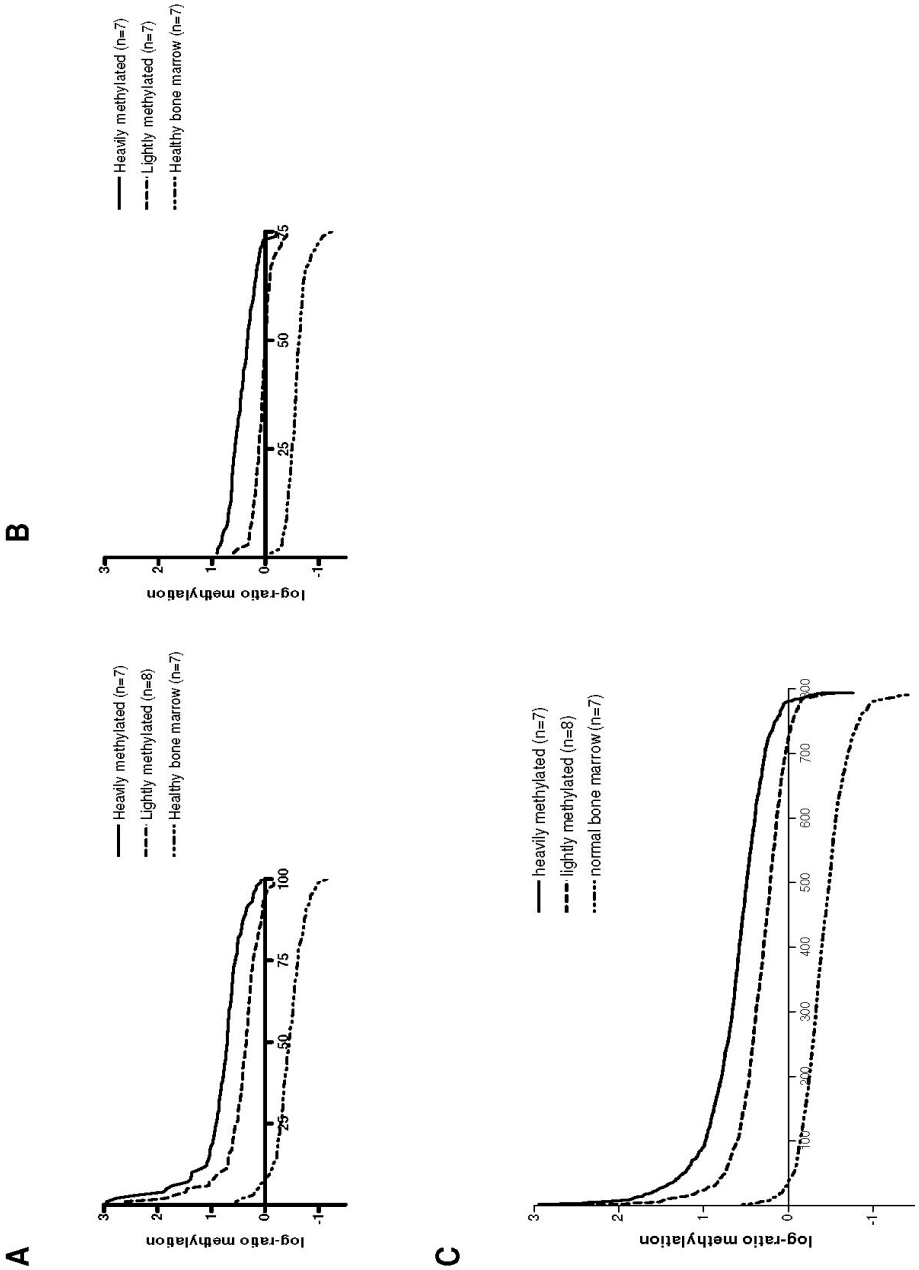
For the one-step Cox model patients were stratified as high-risk when aged <6 months (183 days) and displaying white blood cell counts (WBC) $>300 \times 10^9$/L at diagnosis. Alternatively, patients were classified as medium-risk when aged >6 months at presentation, or aged <6 months with WBC $<300 \times 10^9$ / L.

(Statistical) software

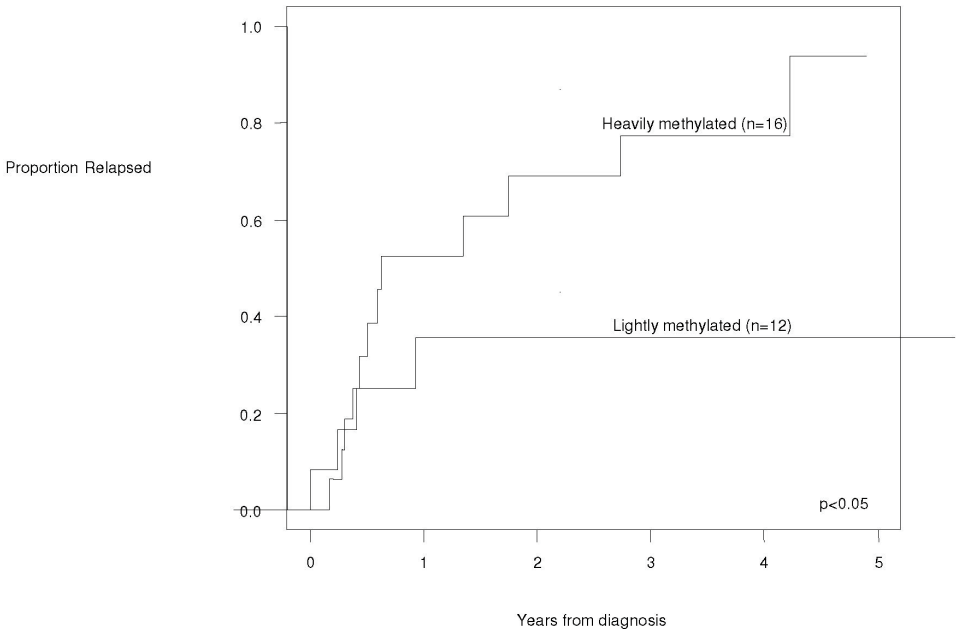
We used the statistical environment R (R Development Core Team, 2007) version 2.6.1 for the microarray analyses, including packages limma,[4] global test,[6] and vsn.[7] Heatmaps were generated in GenePattern version 3.1.1 (Broad Institute, MIT, http://genepattern.broad.mit.edu), and PCA plots were produced using Genemath XT 1.6.1. software (Applied Maths, Inc., Austin TX, USA). SPSS 16.0 statistical software (SPSS Inc., Chicago, IL, USA) was used for computation of survival statistics.



Supplementary Figure S1 - Heatmaps showing the top 20 most significantly hypermethylated probes for each infant ALL subtype (compared with the other subgroups). Additionally, normal bone marrow samples were added to the heatmaps. Columns represent patient samples and rows represent genes. Relative DNA methylation levels are then shown in red (high) and blue (low). Genes and samples were ordered using hierarchical cluster analysis (Euclidean distance, complete linkage) and gene identifiers are listed at the right. Samples are color-coded to indicate the genetic subtype of infant ALL: t(4;11) (red), t(11;19) (orange), t(9;11) (yellow), infant ALL with wild-type *MLL* genes (green) and normal bone marrow (blue). **1-1.** Data are shown for the 9K chip and **1-2**. the 244K chip (Agilent). Due to restricted availability of patient material, Agilent DNA methylation profiles were generated for 49 infant ALL patients (t(4;11)-positive (n=16), t(11;19)-positive (n=15), t(9;11)-positive (n=6), and *MLL* wild-type (n=12)) and 7 normal bone marrows. The 20 most significantly hypermethylated genes shared by t(9;11)-positive and wild-type *MLL* samples were also methylated in normal bone marrow samples, implying that these genes are normally methylated in healthy hematopoietic cells.



Supplementary Figure S2 - Semi-quantitative representation of the degree of methylation. Visualization of the normalized and sorted log-ratios (patient signal divided by common reference signal) of methylation (Y-axis) for the most significantly hypermethylated genes (X-axis) in A. t(4;11) and B. t(11;19)-positive infant ALL patients. Different patient groups represent either heavily or lightly methylated clusters (figure 3). C. Visualization of the normalized and sorted log-ratios of methylation for all significantly hypermethylated genes in t(4;11)-positive infant ALL patients, compared with normal bone marrow samples. These data demonstrate that the observed separation of two patient groups with varying degrees of methylation as observed among both t(4;11) and t(11;19)- positive infant ALL, is not restricted to the most significantly methylated genes. As shown here, these differences remain present when all hypermethylated probes (n=794) are used for t(4;11) positive patients.



Supplementary Figure S3 - Relapse-free survival in t(4;11)-positive and t(11;19)-positive infant ALL patients divided by the degree of DNA methylation (figure 2S), based on hierarchical clustering as shown in figure 3. t(4;11)-positive and t(11;19)-positive patients were combined. Risk of relapse is presented on the Y-axis and the time of follow-up (in years) is presented on the X-axis. The p-value is from a log-rank test. Risk of relapse is significantly increased in the heavily methylated subgroup of infant ALL.

Supplementary Table S1. Patient characteristics. Available at: http://bloodjournal.hematologylibrary.org/content/114/27/5490/suppl/DC1

Supplementary Table S2 - The top 20 of genes most significantly hypermethylated in each genetic subtype of infant ALL (compared with the other subgroups) are listed in order of decreasing statistical significance. Data are shown for the 9K CpG island microarray chip. The Gene ID, Gene Name, log- fold change (logFC), p-value, and p–value adjusted for multiple testing (adj.P.Val) are listed (limma analyses). Results of permutation testing are added for each subgroup (global test). This 9K chip is a custom spotted microarray, therefore no official probe IDs are included. Available at: http://bloodjournal.hematologylibrary.org/content/114/27/5490/suppl/DC1

Supplementary Table S3 - The top 20 of genes most significantly hypermethylated in each genetic subtype of infant ALL (compared with the other subgroups) are listed in order of decreasing statistical significance. Data are shown for the 244K CpG island microarray chip. The probe ID (Agilent ProbeName), Entrez Gene ID, Gene Name, log- fold change (logFC), p–value, and p–value adjusted for multiple testing (adj.P.Val) are shown (limma analyses). Results of permutation testing are added for each subgroup (global test). Available at: http://bloodjournal.hematologylibrary.org/content/114/27/5490/suppl/DC1

Supplementary Table S4. The top 100 of genes most significantly hypermethylated in t(4;11)-positive infant ALL (compared with normal bone marrow) are listed in order of decreasing statistical significance. The probe ID (Agilent ProbeName), Entrez Gene ID, Gene Name, log- fold change (logFC), p–value, and p–value adjusted for multiple testing (adj.P.Val) are shown (limma analyses). The corresponding Affymetrix probe ID from the gene expression platform (Affymetrix ProbeName) is added for these genes as well as the adjusted p-value for differential gene expression (adj.P.Val expression) (limma model). Available at: http://bloodjournal.hematologylibrary.org/content/114/27/5490/suppl/DC1

References

1. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet.* Mar 22 2008;371(9617):1030-1043.
2. Pieters R, Schrappe M, De Lorenzo P, et al. A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet.* Jul 21 2007;370(9583):240-250.
3. Greaves MF. Infant leukaemia biology, aetiology and treatment. *Leukemia.* Feb 1996;10(2):372-377.
4. Jansen MW, Corral L, van der Velden VH, et al. Immunobiological diversity in infant acute lymphoblastic leukemia is related to the occurrence and type of MLL gene rearrangement. *Leukemia.* Apr 2007;21(4):633-641.
5. Hess JL. Mechanisms of transformation by MLL. *Crit Rev Eukaryot Gene Expr.* 2004;14(4):235-254.
6. Pui CH, Gaynon PS, Boyett JM, et al. Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet.* Jun 1 2002;359(9321):1909-1915.
7. Dou Y, Hess JL. Mechanisms of transcriptional regulation by MLL and its disruption in acute leukemia. *Int J Hematol.* Jan 2008;87(1):10-18.
8. Armstrong SA, Staunton JE, Silverman LB, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet.* Jan 2002;30(1):41-47.
9. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell.* Mar 2002;1(2):133-143.
10. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet.* Apr 2001;10(7):687-692.
11. Roman-Gomez J, Castillejo JA, Jimenez A, Barrios M, Heiniger A, Torres A. The role of DNA hypermethylation in the pathogenesis and prognosis of acute lymphoblastic leukemia. *Leuk Lymphoma.* Nov 2003;44(11):1855-1864.
12. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med.* Nov 20 2003;349(21):2042-2054.
13. Stam RW, den Boer ML, Passier MM, et al. Silencing of the tumor suppressor gene FHIT is highly characteristic for MLL gene rearranged infant acute lymphoblastic leukemia. *Leukemia.* Feb 2006;20(2):264-271.
14. Stam RW, den Boer ML, Schneider P, et al. Targeting FLT3 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. *Blood.* Oct 1 2005;106(7):2484-2490.
15. Pocock CF, Malone M, Booth M, et al. BCL-2 expression by leukaemic blasts in a SCID mouse model of biphenotypic leukaemia associated with the t(4;11)(q21;q23) translocation. *Br J Haematol.* Aug 1995;90(4):855-867.
16. Tang R, Faussat AM, Perrot JY, et al. A new acute lymphoblastic leukaemia cell line BEL-1 with t(4; 11) (q21; q23) chromosomal translocation and a unique aberrant p27 transcript. *Br J Haematol.* Sep 2004;126(5):754-755.
17. Stong RC, Korsmeyer SJ, Parkin JL, Arthur DC, Kersey JH. Human acute leukemia cell line with the t(4;11) chromosomal rearrangement exhibits B lineage and monocytic characteristics. *Blood.* Jan 1985;65(1):21-31.
18. Huang TH, Perry MR, Laux DE. Methylation profiling of CpG islands in human breast cancer cells. *Hum Mol Genet.* Mar 1999;8(3):459-470.
19. Yan PS, Wei SH, Huang TH. Differential methylation hybridization using CpG island arrays. *Methods Mol Biol.* 2002;200:87-100.

20. Yoo CB, Cheng JC, Jones PA. Zebularine: a new drug for epigenetic therapy. *Biochem Soc Trans.* Dec 2004;32(Pt 6):910-912.
21. Marquez VE, Barchi JJ, Jr., Kelley JA, et al. Zebularine: a unique molecule for an epigenetically based strategy in cancer chemotherapy. The magic of its chemistry and biology. *Nucleosides Nucleotides Nucleic Acids.* 2005;24(5-7):305-318.
22. Pieters R, Loonen AH, Huismans DR, et al. In vitro drug sensitivity of cells from children with leukemia using the MTT assay with improved culture conditions. *Blood.* Dec 1 1990;76(11):2327-2336.
23. Zahurak M, Parmigiani G, Yu W, et al. Pre-processing Agilent microarray data. *BMC Bioinformatics.* 2007;8:142.
24. Smyth GK. Limma: linear models for microarray data. . In: R. Gentleman VC, S. Dudoit, R. Irizarry, W. Huber, ed. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer; 2005:397-420.
25. Benjamini Y HY. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J Roy Stat Soc B* 1995;57(1):289-300
26. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics.* Jan 1 2004;20(1):93-99.
27. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer.* Nov 2007;7(11):823-833.
28. Fraga MF, Esteller M. Towards the human cancer epigenome: a first draft of histone modifications. *Cell Cycle.* Oct 2005;4(10):1377-1381.
29. Milne TA, Briggs SD, Brock HW, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell.* Nov 2002;10(5):1107-1117.
30. Nakamura T, Mori T, Tada S, et al. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell.* Nov 2002;10(5):1119-1128.
31. Mueller D, Bach C, Zeisig D, et al. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood.* Dec 15 2007;110(13):4445-4454.
32. Campo Dell'Orto M, Banelli B, Giarin E, et al. Down-regulation of DLX3 expression in MLL-AF4 childhood lymphoblastic leukemias is mediated by promoter region hypermethylation. *Oncol Rep.* Aug 2007;18(2):417-423.
33. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet.* Apr 2007;39(4):457-466.
34. Paredes J, Albergaria A, Oliveira JT, Jeronimo C, Milanezi F, Schmitt FC. P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. *Clin Cancer Res.* Aug 15 2005;11(16):5869-5877.
35. Dobrzycka KM, Kang K, Jiang S, et al. Disruption of scaffold attachment factor B1 leads to TBX2 up-regulation, lack of p19ARF induction, lack of senescence, and cell immortalization. *Cancer Res.* Aug 15 2006;66(16):7859-7863.
36. Erfurth FE, Popovic R, Grembecka J, et al. MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. *Proc Natl Acad Sci U S A.* May 27 2008;105(21):7517-7522.
37. Krivtsov AV, Feng Z, Lemieux ME, et al. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell.* Nov 4 2008;14(5):355-368.

Concluding remarks and future perspectives

Chapter 6

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

DNA methylation profiling: different techniques, different targets

The past decade has seen major advances in the methods for detecting DNA methylation on a genome-wide level. Enzyme-based (e.g., HELP[1], MMASS[2], DMH[3] and CHARM[4]) and chromatin immunoprecipitation methods[5] are most commonly used in combination with custom or commercial CpG island or promoter microarrays. Recently, captured methylated DNA[6] and bisulfite-converted reduced representations[7] were analyzed using high-throughput sequencing strategies. Although bisulfite-converted reduced representation still relies on restriction enzymes, the combination of antibody-mediated pull-down techniques with sequencing (ChIP-seq) does not. Combining restriction site-independent techniques with high-throughput sequencing can help to approach a true genome-wide analysis of the methylome. However, these ChIP-seq analyses are not without a bias. An over-representation of high CpG content by these techniques can lead to a false negative result of less dense CpG-rich loci. Interestingly, studies using the enzyme-based technique CHARM have shown that conserved regions up to 2 kb from the promoter (CpG island shores) undergo more cancer-related differential methylation than traditional promoter CpG islands do[8]. Additionally, ChIP-on-chip experiments have shown that CpG islands of lower density (intermediate-CpG islands) also undergo more cancer-related differential methylation[9]. Because differential methylation of these regions is tightly correlated to altered gene expression, the importance of these regions might have been underestimated in previous studies. These findings highlight another bias in microarray methodology for detecting DNA methylation. Following the general focus on promoter CpG island regions, the use of arrays that contain only such sites has led to the aforementioned interesting sites being ignored.

The DMH method was used in **Chapters 2, 4 and 5**. This technique -particularly in combination with the home-spotted 8.5 k array (**Chapter 2**)- can be regarded as a low-resolution approach in comparison with ChIP-seq or set-ups using arrays that contain larger genomic coverage. Despite the low genomic representation on the 8.5-k home-spotted microarray, the represented loci were experimentally selected based on their ability to be methylated *in vitro*[10, 11]. Moreover, pre-selection fragmentation was performed using *Mse*I. Because the *Mse*I recognition site rarely occurs in GC-rich regions, this pre-digestion leaves intact most of the low- to high-density CpG island fragments. These aspects of our experimental set-up allowed the identification of colon tumor-specific methylation of the *PTPRGint1* locus in **Chapter 2.** Additionally, in **Chapters 4 and 5**, we used the DMH method in combination with a commercial 244-k oligonucleotide microarray platform containing broader coverage of the genome. This increase in resolution allowed us to identify *BRAF* mutation-specific promoter methylation of *SMO* and *FOXD3* (**Chapter 4**) and *MLL* translocation-specific patterns of DNA methylation (**Chapter 5**).

DNA methylation and early detection of colon cancer

Chapter 2 described the colon tumor-specific methylation of a low CG-dense CpG island that is located in the first intron of the *PTPRG* gene (*PTPRGint1*). No direct biological implication of *PTPRGint1* methylation on colon tumorigenesis could be given in this thesis despite a loss of CTCF binding to the region. However, in sporadic colon lesions ranging from early adenomas to carcinomas, high levels of *PTPRGint1* methylation were observed (**Chapter 2**). In addition, Lynch syndrome-associated colon lesions contained similar high levels of *PTPRGint1* methylation (**Chapter 2**).

As mentioned in the Introduction, early detection greatly increases the survival rate for patients with colorectal cancer. Given the long asymptomatic preclinical phase, population screening of the general public would greatly increase the early detection rate of sporadic colorectal cancer. The available screening tests that are most commonly used for the detection of colorectal cancer include colonoscopy and fecal occult blood tests. Although colonoscopy is a highly sensitive method, its high costs, invasiveness and risk of complications such as bowel perforation make it less applicable for screening the general population. However, it remains a valuable method for screening high-risk patients in Europe and other views on colonoscopy use pervade in the US. Fecal occult blood testing is relatively simple, and although it has low predictive value, several randomized trials[12] have shown a decrease in the mortality rate of colorectal cancer patients by up to 25%. According to published reports, tests using epi- and genetic biomarkers for screening serum and feces promise to hold high value in the development of accurate, non-invasive screening methods[13, 14].

To date, a number of DNA methylation markers have been tested in both stool and serum samples (for an overview, see Table 2). The addition of even more specific and sensitive markers -and a combination of these markers- will provide a screening method that is more accurate, more cost effective and more comfortable than current screening methods. The possible addition of the *PTPRGint1* locus to the current list of stool and/or serum markers will depend on the successful adaption of the locus to an applicable method for such testing. Currently, the use of (quantitative) methylation-specific PCR protocols is predominantly reported in the literature.

Table 2 – DNA methylation markers in colorectal cancer patient serum/plasma and stool

Gene	Sample material	Patients (percentage methylated)	Healty donors (percentage methylated)
ALX4[81]	Stool	25/30 (83%)	16/52 (30%)
CDH4[82]	Peripheral blood	32/46 (70%)	0/17(0%)
CDKN2A/p16[83]	Stool	Adenomas 9/29 (31%)	3/19 (16%)
CDKN2A/p16[84]	Serum	12/17 (71%)	0/10(0%)
GATA4[85]	Stool	20/28 (71%) Cohort 2: 24/47 (51%)	7/45 (16%) 2/30 (7%)
HIC1[86]	Stool	11/26 (42%) Adenomas 4/13 (31%) Hyperplastic polyps 0/9 (0%)	1/32 (3%)
HLTF[87]	Serum	16/49 (32.7%)	3/41(7%)
HLTF[88]	Serum	22/103(21%)	0/20(0%)
ITGA4[89]	Stool	9/13(69%)	6/28(21%)
MGMT[83]	Stool	Adenomas 14/29 (48%)	5/18 (27%)
MLH1[90]	Peripheral blood	35/262 (13.4%)	
MLH1[87]	Serum	19/49 (39%)	1/41 (2%)
NDRG4[91]	Stool	17/28 (61%) Cohort 2: 25/47 (53%)	7/45 (16%) 0/30 (0%)
NGFR[92]	Plasma	68/133 (51%)	29/179 (16%)
OSMR[93]	Stool	26/69 (38%)	4/81 (5%)
PGR[94]	Stool	18/23 (78%)	8/26 (31%)
RUNX3[84]	Serum	11/17 (65%)	0/10 (0%)
SEPT9[92]	Plasma	92/133 (69%)	25/179 (14%)
SFRP2[94]	Stool	19/23 (83%)	6/26 (26%)
SFRP5[94]	Stool	18/23 (78%)	9/26 (35%)
TPEF/HPP1[92]	Plasma	87/133 (65%)	56/179 (31%)
TPEF/HPP1[88]	Serum	13/103 (13%)	0/20 (0%)
Vimentin[95]	Stool	43/94 (46%)	20/198 (10%)

CTCF binding regulation by *PTPRGint1* methylation

The methylation status of *PTPRGint1* was not found to have a direct effect on the transcription level of the *PTPRG* gene (**Chapter 2**). However, the annotation of *PTPRGint1* as a CTCF-binding locus suggests possible downstream effects on transcription and chromatin modifications through the varied functions of CTCF (**Chapter 2**). Recently, an electrophoretic mobility shift assay was performed on the *PTPRGint1* locus and validated this region as a methylation-dependent CTCF binding region (Barry Pepers, Leiden University Medical Center, personal communication).

The binding of CTCF to a region confers protection against DNA methylation; CTCF remains associated with specific chromosomal regions during mitosis, thereby suggesting a possible role in the maintenance of epigenetic marks throughout cell division[15-18]. Aberrant methylation of CTCF binding sites such as the ones that are associated with *ARF*[19, 20], *Rb*[21], *BRCA1*[22, 23], *p16*,[24] *RASSF1*[24] and *CDH1*[24] result in the spreading of DNA methylation to the promoters of these genes, as implicated previously in human cancers. CTCF can directly protect the binding sequence against DNA methylation; inhibition of CTCF binding -and consequently, its insulator function- putatively results in a boundary loss between euchromatin and heterochromatin with sequential spreading of inactivating epigenetic marks (Figure 8)[24-26].

In humans, the *CTCF* gene maps to the cancer-associated chromosomal locus 16q22.1 (Ref. 27), which is a region that is often lost in primary breast carcinomas[28-30], prostate adenocarcinomas[31], ovarian cancer[32] and Wilms' tumors[33]. However, the biallelic loss of CTCF expression is not observed in human cancers. CTCF null mice exhibit early embryonic lethality; conditional knock-down of CTCF in cultured fibroblasts rapidly leads to apoptosis[34, 35], thereby suggesting that a loss of CTCF is incompatible with cell survival. However, tumor-specific mutations in the third and seventh zinc fingers of the *CTCF* 11-zinc finger coding domain have been found in breast, prostate and Wilms' tumors[36]. All of the identified mutations were accompanied by the loss of the second *CTCF* allele and resulted in a missense codon at a position that is predicted to be critical for either zinc finger formation or DNA sequence recognition[36]. CTCF can use various combinations of zinc fingers to bind to a wide range of DNA sequences and proteins. Theoretically, the mutations that were identified in *CTCF* could confer a differential loss of binding to specific DNA recognition sites and/or proteins.

Such a model wherein specific CTCF functions are affected without interfering with other CTCF functions that are essential for cell viability is potentially interesting. However, mutations in *CTCF* are extremely rare and cannot account for the wide-scale DNA methylation differences that are observed in cancers[33, 35-37]. An alteration in the protein partners that CTCF requires to perform one or more of its many functions is an alternative explanation.

Although the spreading of heterochromatin marks can be a consequence of a loss of CTCF binding, evidence for the sequential progression between CTCF binding and DNA methylation of its own binding site remains elusive. Given that CTCF binding to a CpG-containing binding site has been shown to be sensitive to methylation, whether methylation of the CTCF binding site is caused by a loss of CTCF binding or whether aberrant methylation precedes CTCF binding loss remains an open question.

Before tackling these important fundamental questions, the impact of losing CTCF

binding after methylation of the *PTPRGint1* region must first be investigated. Mapping of CTCF binding sites in normal colon epithelium would serve as an essential validation of the region as a colon-specific CTCF binding site, as the *PTPRGint1* region has only been described as such in a cell line model system (**Chapter 2**). Applying the chromosome conformation capture technique to both normal colon epithelium and colon cancers with *PTPRGint1* methylation might reveal putative differences in chromatin loop formation. An analysis of DNA/histone modifications, Lamina-associated domains and transcription changes with respect to these putative chromatin loop formation alterations might provide a clear view of the consequences of *PTPRGint1* methylation. However, apart from assessing DNA methylation and mRNA levels, the aforementioned analyses are difficult to perform on tissues other than cultured cells due to the high numbers of cells that are required.

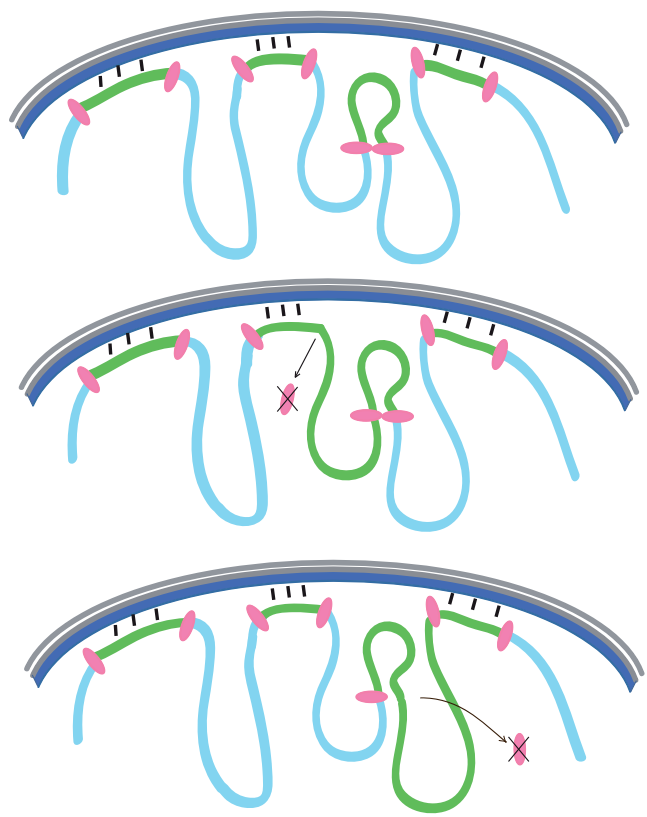


Figure 8 – Heterochromatin spreading caused by disassociation of CTCF binding. Upper panel: normal CTCF protein (pink) binding demarcates the boundaries between heterochromatin (green line) and euchromatin (light blue line). Middle panel: disassociation of CTCF protein (pink) binding causes spreading of lamina-associated heterochromatic marks. Lower panel: spreading of non-lamina-associated heterochromatic marks by the regional loss of CTCF binding. Figure adapted from de Wit et al.[25].

Aberrant histone modification machinery: initiation of cancer specific DNA methylation?

In support of a direct link between histone modifications and DNA methylation, H3K4 methylation was found to protect against DNA methylation[38]. In addition, binding of the PRC2 member EZH2 to a region recruits DNMTs[39]. Evidence for a developmental model (see Figure 4 in the General Introduction) resulted from the discovery that polycomb-mediated H3K27 methylation in ES cells pre-marks genes for *de novo* DNA methylation in colon cancer[40-45]. In this model, the balance between mediator binding of inactivating (Polycomb Repressive Complex 2 ; PRC2) and activating histone (Trithorax-group ; TrxG) marks dictate the downstream DNA methylation and expression states of chromosomal regions (see Figure 4 in the General Introduction). This balance between the antagonistic histone modifications is most pronounced in ES cells in which almost all H3K27^{me3} bound promoters are bivalent[46, 47]. The loss of this bivalent binding in both adult differentiated colon tissue and colon cancer supports the conventional view that such histone modifications act as 'plastic' epigenetic switches for transcription and *de novo* methylation during development[40-45]. In this hypothesis, DNA methylation serves as a definitive epigenetic lock that is preceded by histone modifications and their associated machinery.

Recent studies have reported a possible pre-marking of cancer-specific hypermethylated genes by H3K27^{me3} and binding of the PRC2 member SUZ12 in both ES cells and differentiated normal colon mucosa[42-44]. The presence of repressive histone modifications at promoters in normal colon epithelium suggests that the associated genes are in a transcriptionally silent state prior to tumor formation, thereby reducing the relevance of their promoter's methylation on expression and thus tumorigenesis. To focus on genes that are silenced by tumor-specific promoter methylation (rather than lineage-specific methylation), we excluded from our analysis in Chapter 4 the loci with H3K27^{me3} in embryonic stem (ES) cells.

In model systems, abnormal TrxG or PRC2 function often results in aberrant gene expression that leads to tumor development, which suggests that dysregulation of these epigenetic programs can initiate tumorigenesis[39, 48-50]. Although there is no evidence of alterations in the balance between TrxG proteins and PRC2 in colon cancer, such an alteration exists in infant B-ALL. In Chapter 5, we described the translocation-based grouping of DNA methylation patterns from *MLL*-rearranged B-ALL patients. Although patients who contain the translocation t(4;11) or t(11;19) have comparable high levels of DNA methylation, patients with a t(9;11) translocation have DNA methylation patterns that are comparable to normal controls and B-ALL patients without *MLL* mutations. As the histone methyltransferase moiety of one copy of *MLL* is lost in all three translocation subtypes, their differing DNA methylation levels suggest that this loss has minimal impact on DNA methylation. In addition, 90% of all genes that are expressed in t(4;11) rearranged leukemia cells contain the H3K4^{me3} modification[51]. Therefore, perturbation of MLL H3K4 methyltransferase activity by the alteration of one copy of *MLL* also has a low impact on H3K4 methylation levels[52]. Finally, recent studies have shown that a loss of the normal H3K4 histone methyltransferase activity in one copy of *MLL* is not sufficient to initiate leukemia[53, 54]. It is therefore likely that the combined functions of the reciprocal fusion partners facilitate both aberrant DNA methylation and tumorigenesis in infant B-ALL[54, 55].

The fusion protein that results from the aforementioned translocations is a combination of a C-terminal transcriptional activation domain provided by the fusion partner (AF4, AF9 or ENL) fused with the DNA binding N terminus of MLL. Recruitment of DOT1L and sequential H3K79 methylation to regions that are bound by N-terminal MLL is contributed to these fusion proteins[51, 52, 56-59]. This addition of H3K79^{me} -a modification that is associated with transcription elongation- could explain the high expression of MLL target genes that are associated with leukemogenesis such as *MEIS1* and *HOXA9*[51, 52, 58]. Co-occupancy of both the germline MLL protein and the MLL fusion protein could provide both the activating H3K4^{me3} histone mark and the transcription elongation mark H3K79^{me}. The combination of both histone modifications might be sufficient to enable such aberrant high expression (Figure 9). However, the effect of the fusion proteins on the DNA hypermethylation that has been observed in infant B-ALL patients with t(4;11) and t(11;19) translocations (Chapter 5) remains unknown.

Recent studies have shown that unlike the MLL-AF4 fusion protein, the t(4;11) reciprocal fusion protein AF4-MLL, which consists of N-terminal AF4 and C-terminal MLL, is sufficient to induce leukemogenesis in mouse models[53]. The N-terminal domain of AF4 has been suggested to retain its ability to bind RNA polymerase II while losing the ability to recruit DOT1L. However, the downstream effects of this reciprocal translocation protein and its importance in leukemogenesis warrant further investigation.

Our future understanding of the full impact of *MLL* rearrangements on histone modifications and their role in aberrant DNA methylation will require multi-dimensional studies that combine histone modification, DNA methylation, insulator binding, transcription and MLL fusion protein binding data. The involvement of the TrxG member MLL supports a direct link to the disrupted balance between the antagonistic histone modifiers PRC2 and TrxG in infant B-ALL. However, down-regulation of MLL histone-methyltransferase activity has little effect on H3K4 methylation levels[52]. Complete mapping of DNA and histone modifications -combined with transcription data from the various stages in hematopoiesis and *MLL*-rearranged leukemogenesis- will provide a clearer understanding of the true identity of the oncogenic drivers. The identification of translocation-specific MLL binding sites might provide additional insight into the differences in DNA methylation that have been observed between infant B-ALL patients who harbor different types of translocations (Chapter 5).

A sequential model for colon differentiation and tumorigenesis would provide similar insights and could reveal possible links between aberrant DNA methylation, histone modification and their modifier complexes. Linking this data to transcription and insulator binding data should help expand our understanding of the role of DNA methylation in the etiology of colon cancer.

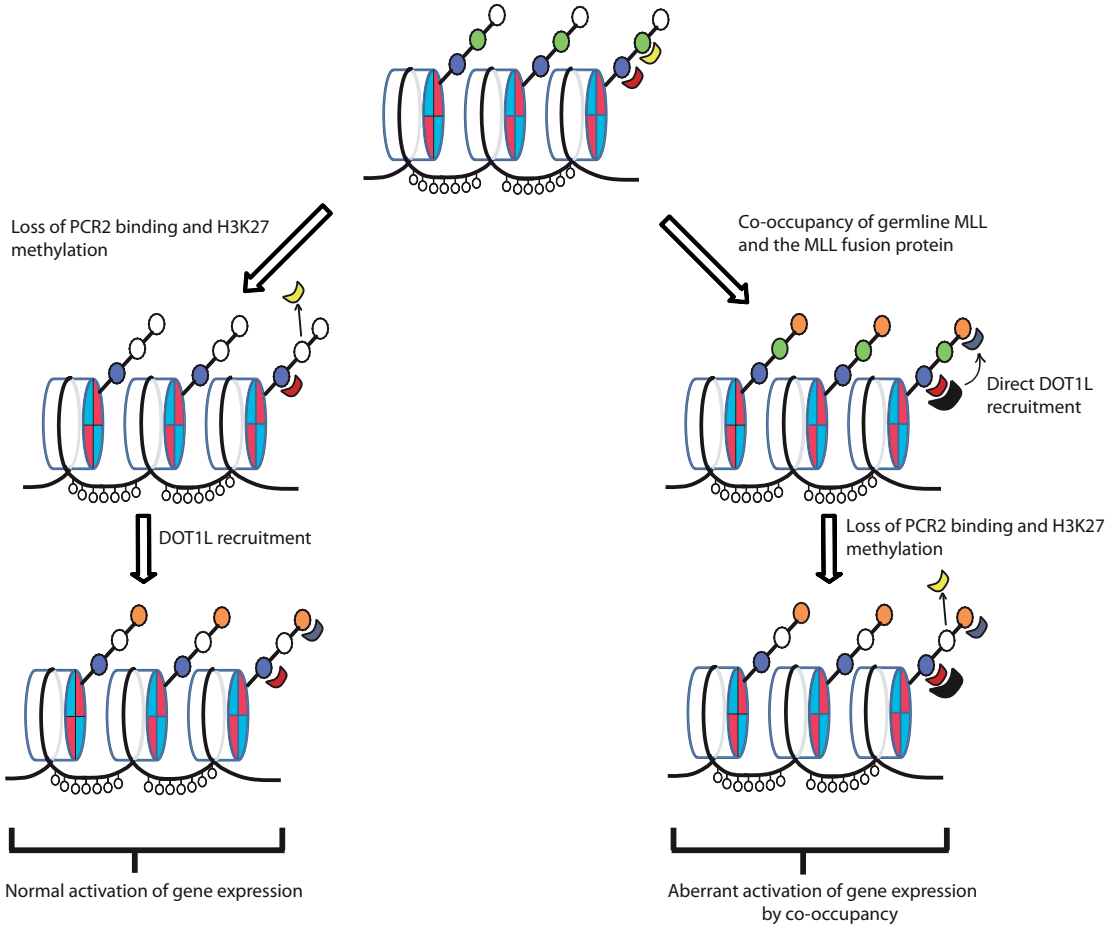


Figure 9 – Progression model showing normal and aberrant recruitment of the H3K79 histone methyltransferase DOT1L. Co-occupancy of both the unaffected MLL (red crescent) as well as the MLL fusion protein (large red crescent with gray stripes) recruits DOT1L (grey crescent). This direct recruitment of DOT1L and its histone (H3K79) mark (orange) for transcription elongation can aberrantly activate gene expression by circumventing the sequential progression and balance between the TrxG/MLL, PRC2 (yellow crescent) proteins and associated histone marks (H3K4 methylation in purple and H3K27 methylation in green).

The TrxG family and BRAF, is there a connection?

The aforementioned enrichment of PRC2 components and H3K27^{me3} in targets for colon cancer-specific DNA methylation suggests a possible dysregulation of the balance between TrxG and PRC2 during colon tissue differentiation (**Chapter 4**). Although this balance is perturbed by deregulation of the TrxG protein MLL in *MLL*-rearranged ALL, such an association has not been described in colon cancer. One study described a direct link between colon cancer and dysregulation of TrxG in both colon cancer cell lines and tumors. The TrxG member *MLL2* was found to be overexpressed in colon tumors compared to corresponding adjacent tissue[60]. In addition, colon cancer cell lines that were derived from highly invasive, poorly differentiated tumors exhibited altered sub-cellular distribution and proteolytic processing of MLL2 compared to non-tumor cell lines

and less invasive tumor cell lines[60]. In contrast to increased *MLL2* expression, the altered sub-cellular distribution and proteolytic processing indicate a decrease in normal MLL2 activity. Although the relationship between CIMP and MLL2 has not been investigated, CIMP is known to be associated with poor differentiation, thus indicating a possible association between CIMP and the proteolytic dysregulation or sub-cellular distribution of MLL2. Unfortunately, no distinction of CIMP, MSI or *BRAF* mutational status was made in the study in relation to the overexpression of *MLL2* in colon tumors and cell lines. Grouping these factors would provide a clearer insight into the role of MLL2 in the etiology of tumors with aberrant DNA methylation. Interestingly, in **Chapter 4** we described an association of colon cancer DNA methylation with H3K27^{me3} targets in ES cells and showed that this enrichment is less pronounced in CIMP- and *BRAF*-associated DNA methylation.

An association between *BRAF* mutations and CIMP colon cancer has been documented[61] and was also described in **Chapter 3** for colon cancer patients up to 50 years of age. CIMP colon cancer with *BRAF* mutations and MSI may originate from sessile serrated polyps through a unique pathway. While *BRAF* mutations are present in sessile serrated polyps and serrated aberrant crypt foci, *KRAS* mutations are more closely associated with non-serrated polyps and non-serrated aberrant crypt foci[62-67]. Interestingly, 90% of all aberrant crypt foci with *BRAF* mutations were found to be microsatellite-stable , whereas ~70% of sporadic MSI colon cancers exhibit *BRAF* mutations. These observations led to the suggestion that *BRAF* mutations precede *MLH1* methylation and that mutationally active BRAF might play an initiating role in the manifestation of aberrant DNA methylation in colon cancer[62-64, 68-70]. Although no direct interactions between members of the RAS-RAF pathway and either the TrxG, PRC2 or DNMT family have been described, the downstream transcription factors that are activated by the RAS-RAF pathway protein MAPK (for example, c-Fos, CREB and c-MYC) interact with DNMT3A and DNMT3B[71]. Additionally, it was shown that inhibiting the ERK/MAPK signaling pathway decreases the genomic DNA methylation content in cancer cells[72]. KRAS and BRAF are directly linked in the same pathway but associate differently with CIMP in colon cancer. Therefore, interactions between these shared transcription factors and the DNA methylation machinery should be investigated further. An alternate hypothesis regarding the association between *BRAF* mutations and DNA methylation suggests that promoter methylation and silencing of specific target genes could favor the selection of activating *BRAF* mutations. This hypothesis has been suggested for *IGFBP7*, a mediator of BRAF-induced cellular senescence[73-75]. Promoter methylation and the down-regulation of *IGFBP7* could provide a favorable context in which to obtain an activating *BRAF* mutation[73-75]. In **Chapter 4,** we showed *BRAF* mutation-specific promoter methylation of *FOXD3* after stringent filtering for pre-marking and copy number changes. FOXD3 is a mediator of p21^{Cip1}- and p53-dependent cell cycle arrest, which is down-regulated by constitutively active BRAF in melanoma cells[76]. These findings suggest that FOXD3 might play a role in eluding BRAF-induced senescence in colon cancer through epigenetic inactivation. The augmented proliferation and high levels of senescence that are induced by constitutively active BRAF could favor the selection of cells that can escape senescence by the epigenetic inactivation of mediators, which might explain the link between *BRAF* mutations and aberrant DNA methylation in colon cancer.

In **Chapter 3**, we described site-specific methylation of the *MLH1* promoter in sporadic colon cancer patients with an age of diagnosis below 50 years, despite low levels of global methylation. In contrast, patients with a *BRAF* mutation who were either below

or above 50 years of age at diagnosis have significantly higher levels of CIMP marker methylation. The aforementioned *BRAF* mutation-driven selection for cells that can escape senescence suggests an accumulation of promoter methylation that ultimately results in MMR deficiency and/or escape from BRAF-induced senescence. Future studies of the interactions of activated RAS-RAF pathway proteins with histone and DNA modification machineries should yield a clearer understanding of the roles of KRAS and BRAF in the initiation of aberrant DNA methylation. The suggested interaction between RAS-RAF activated transcription factors and DNMT3A and DNMT3B in targeting specific promoter loci[71] is particularly interesting. Identifying differential RAS-RAF-associated transcription factor binding sites in *BRAF* and *KRAS* mutated cancer models might point towards specific mediators of *BRAF*-initiated DNA methylation. However, the mechanisms that underlie the observed site-specific *MLH1* promoter methylation in *BRAF* wild-type patients have been elusive. Polymorphisms such as the *MLH1* -93G>A polymorphism that was discussed in **Chapter 3**, when present in such a transcription factor binding site and thereby putatively altering binding, could provide an additional factor in site-specific DNA methylation.

The observation that promoter methylation of the Methylguanine-DNA methyltransferase (*MGMT*) gene can give rise to mutations in *KRAS* and the tumor suppressor gene *TP53*[77] is somewhat contradictory. A loss of *MGMT* expression can result in a G-to-A transition in *TP53* and codons 12 and 13 of the *KRAS* gene[78-81]. These findings argue against a role for the RAS-RAF pathway in initiating aberrant DNA methylation; elucidating the precise cascade of events is required to determine the function of RAS-RAF pathway in this process.

The role of epigenetics in the initiation of colon cancer and its association with *KRAS* and *BRAF* mutations is complex and requires additional research to unravel the molecular mechanisms that are involved. Although links between mutations in the RAS-RAF pathway and the DNA methylation machinery have been reported[68, 71], these links are indirect. In addition, cellular model systems with constitutively active BRAF fail to induce CIMP[73]. Selective pressure on neoplastic cells that carry *KRAS* or *BRAF* mutations for the epigenetic inactivation of certain genes might be a more plausible explanation for the associations between the mutations and the increased levels of promoter DNA methylation in colon cancer. The absence of *BRAF* mutations and the relatively low levels of methylation that are observed in Lynch syndrome patients[82], young sporadic colon cancer patients with site-specific *MLH1* methylation and colon cancer patients with germline *MLH1* methylation (**Chapter 3**) also hint towards such an association.

References

1. Khulan B, Thompson RF, Ye K, Fazzari MJ, Suzuki M, Stasiek E, Figueroa ME, Glass JL, Chen Q, Montagna C, Hatchwell E, Selzer RR, Richmond TA, Green RD, Melnick A, Greally JM. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. Genome Res 2006;16:1046-1055.
2. Ibrahim AE, Thorne NP, Baird K, Barbosa-Morais NL, Tavare S, Collins VP, Wyllie AH, Arends MJ, Brenton JD. MMASS: an optimized array-based method for assessing CpG island methylation. Nucleic Acids Res 2006;34:e136.
3. Huang TH, Perry MR, Laux DE. Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet 1999;8:459-470.
4. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Wen B, Feinberg AP. Comprehensive high-throughput arrays for relative methylation (CHARM). Genome Res 2008;18:780-790.
5. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005;37:853-862.
6. Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM, Thorne NP, Backdahl L, Herberth M, Howe KL, Jackson DK, Miretti MM, Marioni JC, Birney E, Hubbard TJ, Durbin R, Tavare S, Beck S. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 2008;26:779-785.
7. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 2008;454:766-770.
8. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009;41:178-186.
9. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457-466.
10. Yan PS, Chen CM, Shi H, Rahmatpanah F, Wei SH, Caldwell CW, Huang TH. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. Cancer Res 2001;61:8375-8380.
11. Yan PS, Efferth T, Chen HL, Lin J, Rodel F, Fuzesi L, Huang TH. Use of CpG island microarrays to identify colorectal tumors with a high degree of concurrent methylation. Methods 2002;27:162-169.
12. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol 2008;103:1541-1549.
13. Kim MS, Lee J, Sidransky D. DNA methylation markers in colorectal cancer. Cancer Metastasis Rev 2010;29:181-206.
14. Wong IH. Methylation profiling of human cancers in blood: molecular monitoring and prognostication (review). Int J Oncol 2001;19:1319-1324.
15. Burke LJ, Zhang R, Bartkuhn M, Tiwari VK, Tavoosidana G, Kurukuti S, Weth C, Leers J, Galjart N, Ohlsson R, Renkawitz R. CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. EMBO J 2005;24:3291-3300.
16. Caiafa P, Zlatanova J. CCCTC-binding factor meets poly(ADP-ribose) polymerase-1. J Cell Physiol 2009;219:265-270.

17. Guastafierro T, Cecchinelli B, Zampieri M, Reale A, Riggio G, Sthandier O, Zupi G, Calabrese L, Caiafa P. CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. J Biol Chem 2008;283:21873-21880.
18. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A 2008;105:8309-8314.
19. Rodriguez C, Borgel J, Court F, Cathala G, Forne T, Piette J. CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus. Biochem Biophys Res Commun 2010;392:129-134.
20. Tam AS, Devereux TR, Patel AC, Foley JF, Maronpot RR, Massey TE. Perturbations of the Ink4a/Arf gene locus in aflatoxin B1-induced mouse lung tumors. Carcinogenesis 2003;24:121-132.
21. De La Rosa-Velazquez IA, Rincon-Arano H, Benitez-Bribiesca L, Recillas-Targa F. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. Cancer Res 2007;67:2577-2585.
22. Butcher DT, Rodenhiser DI. Epigenetic inactivation of BRCA1 is associated with aberrant expression of CTCF and DNA methyltransferase (DNMT3B) in some sporadic breast tumours. Eur J Cancer 2007;43:210-219.
23. Butcher DT, Mancini-DiNardo DN, Archer TK, Rodenhiser DI. DNA binding sites for putative methylation boundaries in the unmethylated region of the BRCA1 promoter. Int J Cancer 2004;111:669-678.
24. Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. Mol Cell 2009;34:271-284.
25. Tiwari VK, Baylin SB. Breaching the boundaries that safeguard against repression. Mol Cell 2009;34:395-397.
26. Filippova GN, Lindblom A, Meincke LJ, Klenova EM, Neiman PE, Collins SJ, Doggett NA, Lobanenkov VV. A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. Genes Chromosomes Cancer 1998;22:26-36.
27. Cleton-Jansen AM, Moerland EW, Kuipers-Dijkshoorn NJ, Callen DF, Sutherland GR, Hansen B, Devilee P, Cornelisse CJ. At least two different regions are involved in allelic imbalance on chromosome arm 16q in breast cancer. Genes Chromosomes Cancer 1994;9:101-107.
28. Collins C, Volik S, Kowbel D, Ginzinger D, Ylstra B, Cloutier T, Hawkins T, Predki P, Martin C, Wernick M, Kuo WL, Alberts A, Gray JW. Comprehensive genome sequence analysis of a breast cancer amplicon. Genome Res 2001;11:1034-1042.
29. Rakha EA, Armour JA, Pinder SE, Paish CE, Ellis IO. High-resolution analysis of 16q22.1 in breast carcinoma using DNA amplifiable probes (multiplex amplifiable probe hybridization technique) and immunohistochemistry. Int J Cancer 2005;114:720-729.
30. Latil A, Cussenot O, Fournier G, Driouch K, Lidereau R. Loss of heterozygosity at chromosome 16q in prostate adenocarcinoma: identification of three independent regions. Cancer Res 1997;57:1058-1062.
31. Iwabuchi H, Sakamoto M, Sakunaga H, Ma YY, Carcangiu ML, Pinkel D, Yang-Feng TL, Gray JW. Genetic analysis of benign, low-grade, and high-grade ovarian tumors. Cancer Res 1995;55:6172-6180.
32. Yeh A, Wei M, Golub SB, Yamashiro DJ, Murty VV, Tycko B. Chromosome arm 16q in Wilms tumors: unbalanced chromosomal translocations, loss of heterozygosity, and assessment of the CTCF gene. Genes Chromosomes Cancer 2002;35:156-163.

33. Docquier F, Farrar D, D'Arcy V, Chernukhin I, Robinson AF, Loukinov D, Vatolin S, Pack S, Mackay A, Harris RA, Dorricott H, O'Hare MJ, Lobanenkov V, Klenova E. Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. Cancer Res 2005;65:5112-5122.
34. Filippova GN. Genetics and epigenetics of the multifunctional protein CTCF. Curr Top Dev Biol 2008;80:337-360.
35. Filippova GN, Qi CF, Ulmer JE, Moore JM, Ward MD, Hu YJ, Loukinov DI, Pugacheva EM, Klenova EM, Grundy PE, Feinberg AP, Cleton-Jansen AM, Moerland EW, Cornelisse CJ, Suzuki H, Komiya A, Lindblom A, Dorion-Bonnet F, Neiman PE, Morse HC, III, Collins SJ, Lobanenkov VV. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. Cancer Res 2002;62:48-52.
36. Zhou XL, Werelius B, Lindblom A. A screen for germline mutations in the gene encoding CCCTC-binding factor (CTCF) in familial non-BRCA1/BRCA2 breast cancer. Breast Cancer Res 2004;6:R187-R190.
37. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448:714-717.
38. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van EA, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di CL, de LY, Fuks F. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006;439:871-874.
39. Hosey AM, Chaturvedi CP, Brand M. Crosstalk between histone modifications maintains the developmental pattern of gene expression on a tissue-specific locus. Epigenetics 2010;5:273-281.
40. Kondo Y. Epigenetic cross-talk between DNA methylation and histone modifications in human cancers. Yonsei Med J 2009;50:455-463.
41. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007;39:237-242.
42. Rada-Iglesias A, Enroth S, Andersson R, Wanders A, Pahlman L, Komorowski J, Wadelius C. Histone H3 lysine 27 trimethylation in adult differentiated colon associated to cancer DNA hypermethylation. Epigenetics 2009;4:107-113.
43. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 2007;39:232-236.
44. Tiwari VK, McGarvey KM, Licchesi JD, Ohm JE, Herman JG, Schubeler D, Baylin SB. PcG proteins, DNA methylation, and gene repression by chromatin looping. PLoS Biol 2008;6:2911-2927.
45. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007;448:553-560.
46. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, Orlov YL, Sung WK, Shahab A, Kuznetsov VA, Bourque G, Oh S, Ruan Y, Ng HH, Wei CL. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 2007;1:286-298.
47. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001;20:5695-5707.
48. Sparmann A, van LM. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer 2006;6:846-856.

49. Taghavi P, van LM. Developmental biology: two paths to silence merge. Nature 2006;439:794-795.
50. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell 2007;130:77-88.
51. Bursen A, Schwabe K, Ruster B, Henschler R, Ruthardt M, Dingermann T, Marschalek R. The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL. AF4. Blood 2010;115:3570-3579.
52. Collins EC, Pannell R, Simpson EM, Forster A, Rabbitts TH. Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. EMBO Rep 2000;1:127-132.
53. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 2007;7:823-833.
54. Zhang W, Xia X, Reisenauer MR, Hemenway CS, Kone BC. Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCalpha in an aldosterone-sensitive manner. J Biol Chem 2006;281:18059-18068.
55. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y. hDOT1L links histone methylation to leukemogenesis. Cell 2005;121:167-178.
56. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. Cancer Res 2005;65:11367-11374.
57. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia X, Jesneck J, Bracken AP, Silverman LB, Kutok JL, Kung AL, Armstrong SA. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14:355-368.
58. Bitoun E, Oliver PL, Davies KE. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Hum Mol Genet 2007;16:92-106.
59. Natarajan TG, Kallakury BV, Sheehan CE, Bartlett MB, Ganesan N, Preet A, Ross JS, Fitzgerald KT. Epigenetic regulator MLL2 shows altered expression in cancer cell lines and tumors from human breast and colon. Cancer Cell Int 2010;10:13.
60. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-793.
61. Velho S, Moutinho C, Cirnes L, Albuquerque C, Hamelin R, Schmitt F, Carneiro F, Oliveira C, Seruca R. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis? BMC Cancer 2008;8:255.
62. O'Brien MJ. Hyperplastic and serrated polyps of the colorectum. Gastroenterol Clin North Am 2007;36:947-68, viii.
63. O'Brien MJ, Yang S, Mack C, Xu H, Huang CS, Mulcahy E, Amorosino M, Farraye FA. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. Am J Surg Pathol 2006;30:1491-1501.
64. Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, Barker MA, Arnold S, McGivern A, Matsubara N, Tanaka N, Higuchi T, Young J, Jass JR, Leggett BA. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137-1144.
65. Jass JR, Whitehall VL, Young J, Leggett BA. Emerging concepts in colorectal neoplasia. Gastroenterology 2002;123:862-876.

66. Rosenberg DW, Yang S, Pleau DC, Greenspan EJ, Stevens RG, Rajan TV, Heinen CD, Levine J, Zhou Y, O'Brien MJ. Mutations in BRAF and KRAS differentially distinguish serrated versus non-serrated hyperplastic aberrant crypt foci in humans. Cancer Res 2007;67:3551-3554.
67. Ordway JM, Williams K, Curran T. Transcription repression in oncogenic transformation: common targets of epigenetic repression in cells transformed by Fos, Ras or Dnmt1. Oncogene 2004;23:3737-3748.
68. Gazin C, Wajapeyee N, Gobeil S, Virbasius CM, Green MR. An elaborate pathway required for Ras-mediated epigenetic silencing. Nature 2007;449:1073-1077.
69. Nagasaka T, Koi M, Kloor M, Gebert J, Vilkin A, Nishida N, Shin SK, Sasamoto H, Tanaka N, Matsubara N, Boland CR, Goel A. Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950-60, 1960.
70. Hervouet E, Vallette FM, Cartron PF. Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. Epigenetics 2009;4:487-499.
71. Lu R, Wang X, Chen ZF, Sun DF, Tian XQ, Fang JY. Inhibition of the extracellular signal-regulated kinase/mitogen-activated protein kinase pathway decreases DNA methylation in colon cancer cells. J Biol Chem 2007;282:12249-12259.
72. Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, Whitehall VL, Leggett BA, Laird PW. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PLoS One 2009;4:e8357.
73. Wajapeyee N, Serra RW, Zhu X, Mahalingam M, Green MR. Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. Cell 2008;132:363-374.
74. Wajapeyee N, Serra RW, Zhu X, Mahalingam M, Green MR. Role for IGFBP7 in senescence induction by BRAF. Cell 2010;141:746-747.
75. Abel EV, Aplin AE. FOXD3 is a mutant B-RAF-regulated inhibitor of G(1)-S progression in melanoma cells. Cancer Res 2010;70:2891-2900.
76. Patra SK. Ras regulation of DNA-methylation and cancer. Exp Cell Res 2008;314:1193-1201.
77. Abdel-Fattah R, Glick A, Rehman I, Maiberger P, Hennings H. Methylation of the O6-methylguanine-DNA methyltransferase promoter suppresses expression in mouse skin tumors and varies with the tumor induction protocol. Int J Cancer 2006;118:527-531.
78. Esteller M, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Watkins DN, Issa JP, Sidransky D, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res 2000;60:2368-2371.
79. Esteller M, Herman JG. Generating mutations but providing chemosensitivity: the role of O6-methylguanine DNA methyltransferase in human cancer. Oncogene 2004;23:1-8.
80. Wolf P, Hu YC, Doffek K, Sidransky D, Ahrendt SA. O(6)-Methylguanine-DNA methyltransferase promoter hypermethylation shifts the p53 mutational spectrum in non-small cell lung cancer. Cancer Res 2001;61:8113-8117.
81. Ebert MP, Model F, Mooney S, Hale K, Lograsso J, Tonnes-Priddy L, Hoffmann J, Csepregi A, Rocken C, Molnar B, Schulz HU, Malfertheiner P, Lofton-Day C. Aristaless-like homeobox-4 gene methylation is a potential marker for colorectal adenocarcinomas. Gastroenterology 2006;131:1418-1430.
82. Miotto E, Sabbioni S, Veronese A, Calin GA, Gullini S, Liboni A, Gramantieri L, Bolondi L, Ferrazzi E, Gafa R, Lanza G, Negrini M. Frequent aberrant methylation of the CDH4 gene promoter in human colorectal and gastric cancer. Cancer Res 2004;64:8156-8159.
83. Petko Z, Ghiassi M, Shuber A, Gorham J, Smalley W, Washington MK, Schultenover S, Gautam S, Markowitz SD, Grady WM. Aberrantly methylated CDKN2A, MGMT, and MLH1 in colon polyps and in fecal DNA from patients with colorectal polyps. Clin Cancer Res 2005;11:1203-1209.

84. Tan SH, Ida H, Lau QC, Goh BC, Chieng WS, Loh M, Ito Y. Detection of promoter hypermethylation in serum samples of cancer patients by methylation-specific polymerase chain reaction for tumour suppressor genes including RUNX3. Oncol Rep 2007;18:1225-1230.
85. Hellebrekers DM, Lentjes MH, van den Bosch SM, Melotte V, Wouters KA, Daenen KL, Smits KM, Akiyama Y, Yuasa Y, Sanduleanu S, Khalid-de Bakker CA, Jonkers D, Weijenberg MP, Louwagie J, van CW, Carvalho B, Meijer GA, Baylin SB, Herman JG, de Bruine AP, van EM. GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. Clin Cancer Res 2009;15:3990-3997.
86. Lenhard K, Bommer GT, Asutay S, Schauer R, Brabletz T, Goke B, Lamerz R, Kolligs FT. Analysis of promoter methylation in stool: a novel method for the detection of colorectal cancer. Clin Gastroenterol Hepatol 2005;3:142-149.
87. Leung WK, To KF, Man EP, Chan MW, Bai AH, Hui AJ, Chan FK, Sung JJ. Quantitative detection of promoter hypermethylation in multiple genes in the serum of patients with colorectal cancer. Am J Gastroenterol 2005;100:2274-2279.
88. Wallner M, Herbst A, Behrens A, Crispin A, Stieber P, Goke B, Lamerz R, Kolligs FT. Methylation of serum DNA is an independent prognostic marker in colorectal cancer. Clin Cancer Res 2006;12:7347-7352.
89. Ausch C, Kim YH, Tsuchiya KD, Dzieciatkowski S, Washington MK, Paraskeva C, Radich J, Grady WM. Comparative analysis of PCR-based biomarker assay methods for colorectal polyp detection from fecal DNA. Clin Chem 2009;55:1559-1563.
90. Jensen LH, Lindebjerg J, Byriel L, Kolvraa S, Cruger DG. Strategy in clinical practice for classification of unselected colorectal tumours based on mismatch repair deficiency. Colorectal Dis 2008;10:490-497.
91. Melotte V, Lentjes MH, van den Bosch SM, Hellebrekers DM, de Hoon JP, Wouters KA, Daenen KL, Partouns-Hendriks IE, Stessels F, Louwagie J, Smits KM, Weijenberg MP, Sanduleanu S, Khalid-de Bakker CA, Oort FA, Meijer GA, Jonkers DM, Herman JG, de Bruine AP, van EM. N-Myc downstream-regulated gene 4 (NDRG4): a candidate tumor suppressor gene and potential biomarker for colorectal cancer. J Natl Cancer Inst 2009;101:916-927.
92. Lofton-Day C, Model F, Devos T, Tetzner R, Distler J, Schuster M, Song X, Lesche R, Liebenberg V, Ebert M, Molnar B, Grutzmann R, Pilarsky C, Sledziewski A. DNA methylation biomarkers for blood-based colorectal cancer screening. Clin Chem 2008;54:414-423.
93. Kim MS, Louwagie J, Carvalho B, Terhaar Sive Droste JS, Park HL, Chae YK, Yamashita K, Liu J, Ostrow KL, Ling S, Guerrero-Preston R, Demokan S, Yalniz Z, Dalay N, Meijer GA, van CW, Sidransky D. Promoter DNA methylation of oncostatin m receptor-beta as a novel diagnostic and therapeutic marker in colon cancer. PLoS One 2009;4:e6555.
94. Muller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, Muhlthaler M, Ofner D, Margreiter R, Widschwendter M. Methylation changes in faecal DNA: a marker for colorectal cancer screening? Lancet 2004;363:1283-1285.
95. Chen WD, Han ZJ, Skoletsky J, Olson J, Sah J, Myeroff L, Platzer P, Lu S, Dawson D, Willis J, Pretlow TP, Lutterbaugh J, Kasturi L, Willson JK, Rao JS, Shuber A, Markowitz SD. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. J Natl Cancer Inst 2005;97:1124-1132.

Summary
Nederlandse samenvatting
Curriculum Vitae
List of publications

Chapter 7

SUMMARY

The aim of the research that is described in this thesis was to identify DNA methylation that is specific to colon cancer and infant B cell acute lymphoblastic leukemia (B-ALL). To obtain these data, the differential methylation hybridization (DMH) technique was used. In addition, the association between DNA mutations and high levels of cancer-specific methylation was explored.

Chapter 1 provides a general introduction to epigenetics. The complex interplay between the various aspects of DNA methylation, histone modification, nuclear position and chromatin condensation is described. In cancer, aberrant hypomethylation of the genome is accompanied by regional hypermethylation of dense CpG dinucleotide clusters called CpG islands (CGIs). Hypermethylation of promoter CGIs can lead to transcriptional inactivation of the associated gene. This form of epigenetic down-regulation occurs for numerous tumor-suppressor genes in various types of cancer. In this study, the following two types of cancer that undergo cancer-specific DNA methylation changes were investigated: colorectal cancer (CRC) and infant acute lymphoblastic leukemia (ALL) with rearrangements involving the mixed-lineage leukemia (*MLL*) gene.

Chapter 1 also provides a brief overview of CRC tumorigenesis. CRC is the third-most common type of cancer in males and the second-most common in females. The lifetime risk for developing CRC for both men and women is 6%, representing approximately one in 17 individuals within the Netherlands. CRC is one of the leading causes of cancer-related deaths in both Europe and the United States. Approximately half of all CRC patients develop either a localized recurrence or a distant metastasis during the course of their illness, and this dramatically decreases their chance of survival. Therefore, the ability to detect and treat CRC before metastasis occurs is extremely important. A recent decrease in the numbers of new CRC cases and reported deaths has been observed in developed countries, and this is possibly due to improved screening and a consequent increase in early diagnosis. **Chapter 1** also describes the various paths of CRC tumorigenesis and discusses how DNA methylation might affect these pathways.

The second type of cancer that was studied with respect to DNA methylation is infant ALL. Infant patients (i.e., up to one year of age at diagnosis) with *MLL*-rearranged ALL only achieve a 5-year survival rate of approximately 50%. Approximately 80% of infants with ALL carry chromosomal translocations that involve the *MLL* gene, and these patients typically display an immature CD10-negative precursor B-lineage immunophenotype. The *MLL* gene is ubiquitously expressed both during development and in most adult tissues, and its expression is required for definitive hematopoiesis. **Chapter 1** provides both a brief general overview of infant ALL and the specific molecular implications of *MLL* rearrangements in ALL leukemogenesis.

In **Chapter 2**, the differential methylation hybridization (DMH) technique was used in combination with home-spotted arrays to measure DNA methylation in a set of 18 right-sided colon tumor samples. The DMH technique uses three rounds of restriction enzyme digestion of genomic DNA as follows: the first round shears the DNA into CpG-

rich fragments, followed by two methylation-sensitive digestions. These three digestion rounds generate fragments that -upon amplification- can be hybridized to a microarray that contains DNA probes of interest. The microarrays that were used in this chapter were home-spotted arrays that were based on a CpG island clone library that was originally generated at the Sanger Centre from affinity-purified *in vitro* methylated DNA fragments. In this initial setup, we identified colon cancer-specific methylation of a CpG island in the first intron of the *PTPRG* gene (*PTPRGint1*). This methylation pattern on *PTPRGint1* was confirmed by methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) in a larger cohort consisting of the following samples: 18 sporadic adenomas, 67 sporadic carcinomas, 63 Lynch syndrome-associated carcinomas and 119 corresponding normal colon tissue samples. High sensitivity (methylated tumor/total tumors) and high specificity (unmethylated normal/total normal samples) for colorectal cancer was observed during the assessment of *PTPRGint1* methylation, which prompted us to propose the incorporation of this region into new and existing colon cancer screening panels.

Methylation of *PTPRGint1* did not influence the expression of the *PTPRG* gene. However, we found that the *PTPRGint1* region is located in a CTCF binding site. CTCF has been linked to many nuclear mechanisms, including transcriptional regulation, insulation, chromatin condensation and chromosomal loop formation. DNA methylation had a negative effect on the binding of CTCF to this region, which suggests that *PTPRGint1* methylation could alter chromatin density and/or conformation. Such changes in chromatin state could affect the expression of -thus far- unknown genes, thereby affecting colon tumorigenesis.

Methylation of the *MLH1* promoter leads to microsatellite instable (MSI-H) colon cancer which is generally associated with a high age of onset. In **Chapter 3**, a cohort of relatively young colon cancer patients (under 50 years of age at diagnosis) with sporadic *MLH1* methylation was studied. The mutational status of *GADD45A*, *BRAF* and *KRAS*, as well as the presence of the *MLH1* -93G>A polymorphism and germline *MLH1* methylation, were investigated in an attempt to identify initiating factors for the observed DNA methylation. Two individuals exhibited germline *MLH1* methylation; however, locus-specific somatic *MLH1* hypermethylation explained the majority of the sporadic early-onset MSI-H colon cancer cases in our cohort. In this study, we were unable to identify an intrinsic tendency towards CpG island hypermethylation other than aberrant accumulation of CpG island methylation via a somatic mutation of *BRAF*. The site-specific *MLH1* promoter methylation observed in the early-onset MSI-H tumors hints towards an unknown targeted methylation mechanism other than that which is associated with *BRAF* mutations. This hypothesis is discussed further in **Chapter 6**.

The relationship between *BRAF* mutations and DNA methylation was explored further in **Chapter 4**, which describes an improved study of DNA methylation in *BRAF* mutation-associated colon cancer. The DMH technique that was introduced in **Chapter 2** was applied to an oligonucleotide microarray platform that provided broader coverage of genomic CpG islands. To focus on genes that are silenced in a tumor-specific rather than a lineage-specific manner, we excluded regions that exhibit both colon cancer-specific promoter methylation and the so-called histone pre-marking in ES cells. We describe

BRAF mutation-specific promoter methylation of the FOX transcription factor genes *FOXB1, FOXB2* and *FOXD3* and speculate that this methylation might help these tumors escape BRAF-induced senescence.

Chapter 5 describes a study in which the DMH technique was combined with oligonucleotide microarrays that contained high CpG island coverage to determine the methylation patterns of infant B-ALL patients. The majority of *MLL*-rearranged infant ALL cases (i.e., those who are characterized by a t(4;11) or t(11;19) translocation) represent hypermethylated leukemias. In contrast, infant ALL patients with a t(9;11) translocation and those without any *MLL* translocation (wild-type *MLL*) displayed DNA methylation patterns that closely resembled the pattern seen in normal bone marrow. In addition, distinct leukemia-specific DNA methylation patterns were identified in *MLL*-rearranged infant ALL subtypes that were defined by the type of *MLL* translocation or by an absence of such translocations. The majority (90-95%) of the most significantly hypermethylated genes in the t(4;11)- and t(11;19)-positive infant ALL patients were found to be down-regulated as measured at the mRNA level. The high levels of methylation -and its correlation with expression- suggest an epigenetic block of B cell differentiation in infant ALL patients with a t(4;11) or t(11;19) translocation. In addition, the t(4;11)-positive B-ALL cell lines were significantly more sensitive than other leukemia cell lines to the demethylating agent zebularine. This study indicates that patients with a t(4;11) or t(11;19) translocation who have high levels of DNA methylation might be promising candidates for therapies that inhibit DNA methylation.

Chapter 6 presents concluding remarks and provides implications for the future. The collective data that are presented in this thesis suggest that a non-invasive DNA methylation-based screening method using a combination of existing and novel DNA methylation markers (such as *PTPRGint1*) might be feasible for most types of colon cancers. In addition, the translocation-specific DNA methylation that is described in this thesis creates new possibilities for the screening and treatment of aberrant DNA methylation in infant ALL patients with an *MLL* translocation. In **Chapter 6**, we discuss the possible downstream effects of losing CTCF binding after methylation of the *PTPRGint1* region. Known published examples of CTCF binding loss are discussed, and future experiments that aim to gain further insight into the full effects of *PTPRGint1* methylation are proposed.

A link between mutations other than those that occur in *BRAF* and aberrant DNA methylation in CRC was not found in this study. The associations between a constitutively active RAS-RAF pathway and the epigenetics that are described in the literature are discussed in **Chapter 6**. The interplay between *BRAF* mutations and the hypermethylation of the promoters of mediators of oncogene-induced senescence is proposed and discussed. However, the mechanism that underlies the accumulation of aberrant promoter methylation -as can occur in *BRAF*-mutated colon cancer- remains unknown and merits further investigation.

NEDERLANDSE SAMENVATTING

Dit proefschrift beschrijft het onderzoek naar veranderingen in de systemen die de werkzaamheid van ons DNA reguleren. Specifiek werd de rol van de zogenaamde epigenetische DNA veranderingen in dikkedarmkanker en acute lymfatische leukemie (ALL) onderzocht. Dit laatste bij kinderen jonger dan één jaar oud bij diagnose.

Epigenetica is een verzamelnaam voor de verschillende mechanismen die invloed kunnen hebben op het tot expressie komen van onze genen zonder een directe verandering van de DNA code. De best bestudeerde onderdelen van de epigenetica zijn "DNA methylatie" en "covalente modificaties van de histonen".

De menselijke DNA code is samengesteld uit 4 basen: Cytosine, Guanine, Adenine en Thymine. Een cytosine die voorgaat aan een guanine (een CpG-dinucleotide) kan een methylgroep binden aan het vijfde koolstofatoom waardoor er een 5-methyl-cytosine ontstaat. Deze CpG-dinucleotiden zijn vrij zeldzaam in het menselijke genoom en worden veelal in clusters gevonden. Deze clusters worden CpG eilanden genoemd en liggen vlak voor ongeveer 50% van onze genen. De methylatie status van deze CpG-dinucleotide clusters komen in veel gevallen overeen met de transcriptie activiteit van het geassocieerde gen waardoor deze CpG eilanden ervan worden verdacht een rol te spelen in de regulatie van de desbetreffende genen. Dit is het geval bij verscheidene genen die een associatie met kanker hebben. De twee soorten kanker die bestudeerd zijn in dit proefschrift laten veel DNA methylatie verschillen zien ten opzichte van het weefsel waaruit ze ontstaan.

In **Hoofdstuk 1** wordt er een overzicht gegeven over dikkedarmkanker. Dikkedarmkanker is de, op twee na, meest voorkomende soort kanker bij mannen en, op een na, bij vrouwen. De kans om deze kanker te ontwikkelen is ongeveer 6%, of 1 op 17, geldend voor beide seksen in Nederland. Het is een van de meest voorkomende oorzaken van kanker-gerelateerde morbiditeit in Europa en de Verenigde Staten. Bij ongeveer de helft van alle dikkedarmkanker patiënten treedt er een plaatselijke terugkeer van de tumor of een uitzaaiing elders op hetgeen de kans op overlijden doet toenemen. Vroege detectie en behandeling van dikkedarmkanker en haar voorstadia zal de overlevingskans sterk verbeteren. Invoering van presymptomatische screening zou hierin een belangrijke stap kunnen zijn. Ook het gebruik van individuele risicoprofielen kan hierin een belangrijke stap zijn. Te denken valt aan het herkennen van erfelijkheidsfactoren en risicovolle levensstijlen. De moleculaire pathogenese van dikkedarmkanker en hoe DNA methylatie hier een rol in speelt wordt in **Hoofdstuk 1** samengevat.

De tweede soort kanker besproken in dit proefschrift is acute lymfatische leukemie bij kinderen onder de één jaar oud. Deze patiëntengroep heeft, ten opzichte van oudere kinderen met deze ziekte, een opvallend slechte 5-jaar overlevingskans, namelijk 50%. Bij ongeveer 80% van deze kinderen wordt de ziekte veroorzaakt door een uitwisseling tussen chromosomen waarbij het *MLL* gen (gelegen op chromosoom 11), is aangedaan. Dit *MLL* gen is actief in elke cel van ons lichaam en is vooral noodzakelijk voor de ontwikkeling van de bloedcellen. In **Hoofdstuk 1** wordt een kort overzicht van ALL en de moleculaire implicaties van de *MLL* translocaties gegeven.

In **Hoofdstuk 2** wordt de differential methylation hybridization techniek (DMH) toegepast op dikkedarmkanker en normaal darmweefsel. Hiermee konden de DNA methylatie veranderingen in dikkedarmtumoren worden bestudeerd. Deze techniek is gebaseerd op fragmentatie van het DNA met behulp van enzymen die het DNA op specifieke sequenties doorknippen (endonucleases, of restrictie enzymen). Er zijn drie ronden van digestie nodig: één digestie ronde om het DNA in CG-rijke fragmenten te knippen, gevolgd door twee rondes van methylatie-gevoelige digestie. De methylatie-gevoelige restrictie enzymen knippen de ongemethyleerde fragmenten kapot waardoor fragmenten met variërende methylatie, na hybridisatie op een DNA microarray, kunnen worden geïdentificeerd.
Met behulp van deze DMH techniek toonden we dikkedarmkanker-specifieke methylatie aan van een gebied in het eerste intron van het *PTPRG* gen (*PTPRGint1*). Deze specifieke methylatie werd ook gevonden met een tweede techniek (methylation-specific ligation-dependent probe amplification) in een groter cohort van 18 sporadische adenomas, 67 sporadische carcinomas, 63 Lynch syndroom geassocieerde carcinomas en 119 normaal darm weefsel samples.
Methylatie van *PTPRGint1* had geen invloed op de expressie van het *PTPRG* gen. Echter, *PTPRGint1* bleek een zogenaamd methylatiegevoelig CTCF bindingsgebied te zijn, als onderdeel van een mechanisme dat de spatiële organisatie van het DNA reguleert . Verstoring van CTCF binding kan grote implicaties hebben op het expressie patroon door middel van veranderingen in de DNA lusvorming of DNA chromatine status. De methylatie van *PTPRGint1* kan dus, tot nog toe, onbekende gevolgen hebben op de tumorigenese van dikkedarmkanker.

Een subtype van dikkedarmkanker toont zogenaamde microsatelliet instabiliteit (MSI). Hierbij worden repetitieve stukken DNA sequentie (microsatellieten) korter of langer (instabiliteit) door defecten in het DNA mismatch repair systeem. Dit fenotype van dikkedarmkanker wordt gevonden in een erfelijke context (het Lynch syndroom) en in 15 % van niet erfelijke of sporadische darmkanker. Het laatste wordt veroorzaakt door promotor hypermethylatie van het *MLH1* gen en komt vooral voor bij oudere patiënten. In **Hoofdstuk 3** wordt een cohort bestaande uit relatief jonge patiënten met sporadisch MSI dikkedarmkanker beschreven. Verscheidene genetische aberraties (*GADD45A*, *BRAF* en *KRAS* mutaties alsmede het *MLH1* -93G>A polymorfisme) en kiembaan methylatie van *MLH1* werden bestudeerd bij deze patiënten in een poging de initiërende factoren van de *MLH1* hypermethylatie te identificeren. Kiembaan methylering werd gevonden bij twee patiënten. Veelal was er een relatie met het voorkomen van de activerende *BRAF* mutatie met ook hypermethylering van andere target DNA gebieden. In andere gevallen was er sprake van een nog onverklaarde locus-specifieke hypermethylatie van *MLH1*.

De mogelijke relatie tussen *BRAF* mutaties en verhoogde DNA methylatie in kanker wordt verder behandeld in **Hoofdstuk 4.** In dit hoofdstuk wordt de DNA methylatie in *BRAF* gemuteerde dikkedarm tumoren beschreven. De DMH techniek, die ook in **Hoofdstuk 2** wordt beschreven, werd toegepast op een DNA oligonucleotide microarray platform met een hogere genomische CpG eiland dekking. Om ons te concentreren op DNA methylatie met een functionele impact op dikkedarm-tumorigenese hebben we microarray probes met zogenaamde "histon pre-marking" in embryonale stam cellen geëxcludeerd. Kanker-specifieke methylatie van promotoren met deze histon pre-marking lijkt geen tot weinig

effect te hebben op de expressie van een geassocieerd gen. Promotor hypermethylatie zonder histon pre-marking werd aangetoond bij de FOX transcriptie factoren *FOXB1, FOXB2* en *FOXD3*. Deze transcriptie factoren zouden een rol kunnen spelen in het blokkeren van BRAF gedreven oncogenese. De epigenetische uitschakeling van deze factoren zou een manier van de woekerende cel kunnen zijn om deze blokkade te omzeilen.

Hoofdstuk 5 beschrijft het gebruik van de DMH techniek in combinatie met het DNA oligonucleotide microarray platform uit **Hoofdstuk 4** om de methylatiepatronen van kinderen jonger dan een jaar oud met ALL in kaart te brengen. Een meerderheid van de patiënten met een *MLL*-translocatie, vooral met t(4;11) en t(11;19), liet een sterke verhoging van CpG eiland methylatie zien. Een minderheid, vooral met t(9;11), liet echter CpG eiland methylatie niveaus zien die vergelijkbaar zijn met die van normaal beenmerg. Een verdere stratificatie van de patiënten op basis van hun *MLL*-translocatie leidde tot de identificatie van methylatie patronen specifiek voor deze verschillende translocaties. Een merendeel (90-95%) van de gehypermethyleerde genen bij patiënten met een t(4;11) en t(11;19) translocatie kwamen lager tot expressie. De hoge niveaus van methylatie en correlatie met expressie suggereerden een epigenetische blokkade van B-cel differentiatie bij patiënten met een t(4;11) en t(11;19) translocatie. Cellijnen met t(4;11) lieten ten opzichte van leukemie cellijnen zonder deze translocatie, een verhoogde mate van gevoeligheid zien voor het demethylerende middel zebularine. Deze studie toonde aan dat jonge kinderen met ALL en een t(4;11) en t(11;19) translocatie mogelijke kandidaten zijn voor behandeling met DNA methylatie inhiberende therapieën.

Hoofdstuk 6 bevat de conclusies van dit proefschrift en beschrijft mogelijke implicaties voor de toekomst. In **dit hoofdstuk** worden verder de mogelijke gevolgen van *PTPRGint1* methylatie besproken, vooral het verlies van CTCF binding in dit gebied. Bekende gevallen van verlies van CTCF binding uit de literatuur alsmede de benodigde experimenten om een beter inzicht te krijgen in de functie van *PTPRGint1* worden besproken. De in de literatuur beschreven associatie tussen de constante activatie van de RAS-RAF signaalcascade door *BRAF* mutaties en epigenetica wordt verder uitgediept.

CURRICULUM VITAE

Eddy Herman Jasper van Roon werd geboren op 15 februari 1979 te Alphen aan den Rijn. In 1996 behaalde hij het HAVO diploma op het Ashram College in Alphen aan den Rijn en in 1998 behaalde hij het VWO diploma op het Gouwe College in Gouda. Hierna begon hij aan de studie medische biologie aan de Vrije Universiteit (VU) te Amsterdam. Voor deze studie doorliep hij tweemaal een onderzoeksstage. De eerste stage was bij de afdeling Antropogenetica (nu Klinische Genetica) van het "VU Medisch Centrum" in Amsterdam waar hij onder leiding van Dr. H.J. van de Vrugt en Dr. F. Arwert onderzoek deed naar Fanconi anemie. Tijdens zijn tweede stage zocht hij naar kandidaatgenen op chromosoom 16q, die betrokken konden zijn bij borstkanker tumorvorming. Deze stage werd gelopen bij de afdeling Pathologie van het "Leids Universitair Medisch Centrum" (LUMC) in Leiden. In november 2003 werd zijn studie afgerond.

Na zijn afstuderen werkte hij van maart tot oktober 2004 als Onderzoeksassistent bij de afdelingen Pathologie en Humane Genetica van het LUMC. Dit omvatte het onderzoek naar genetische factoren die betrokken zijn bij het ontstaan van tumoren aan de bijschildklier onder leiding van Dr. C.J. Haven en Prof. Dr. H. Morreau.

Van oktober 2004 tot maart 2009 was hij werkzaam als AIO/promovendus bij de afdelingen Pathologie en Humane Genetica van het LUMC onder leiding van Dr. J.M. Boer en Prof. Dr. H. Morreau. De resultaten van dit onderzoek staan beschreven in dit proefschrift. Vanaf april 2009 tot en met maart 2010 was hij tijdelijk werkzaam als postdoctoraal onderzoeker bij de afdeling Pathologie van het LUMC waar hij de CTCF bindingsgebieden in normaal dikkedarmslijmvlies onderzocht. Per april 2010 werkt hij als postdoctoraal onderzoeker bij de afdeling Kindergeneeskunde van het Erasmus Medisch Centrum in Rotterdam.

LIST OF PUBLICATIONS

van Wezel T, Lombaerts M, van Roon EH, Philippo K, Baelde HJ, Szuhai K, Cornelisse CJ, Cleton-Jansen AM.
Expression analysis of candidate breast tumour suppressor genes on chromosome 16q.
Breast Cancer Res. 2005;7(6):R998-1004

Stumpel DJ, van Roon EH, Schneider P Boer JM, de Lorenzo P, Valsecchi MG, de Menezes RX, Pieters R, Stam RW.
Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options.
Blood. 2009 Dec 24;114(27):5490-8

van Roon EH, van Puijenbroek M, Middeldorp A, van Eijk R, de Meijer EJ, Erasmus D, Wouters KA, van Engeland M, Oosting J, Hes FJ, Tops CM, van Wezel T, Boer JM, Morreau H.
Early onset MSI-H colon cancer with MLH1 promoter methylation, is there a genetic predisposition?
BMC cancer 2010, 5;10:180

van Roon EH, de Miranda NF, van Nieuwenhuizen MP, de Meijer EJ, van Puijenbroek M, Yan PS, Huang TH, van Wezel T, Morreau H, Boer JM.
Tumour-specific methylation of *PTPRG* intron 1 locus in sporadic and Lynch syndrome colorectal cancer.
Eur J Hum Genet. 2011 Mar;19(3):307-12