Novel genetic and epigenetic alterations in colorectal tumors and their potential as biomarkers

Terje Cruickshank Ahlquist

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

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

The Research Council of Norway

Norwegian Centre of Excellence
University of Oslo

A thesis for the PhD degree, Oslo, 2008

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 766

ISBN 978-82-8072-325-3

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Unipub AS is owned by
The University Foundation for Student Life (SiO)

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The present work has been performed at the Department of Cancer Prevention, Institute for Cancer Research at Rikshospitalet University Hospital – Radiumhospitalet and was supported by the Research Council of Norway and Centre for Cancer Biomedicine.

I am grateful for working under my supervisor, Professor Ragnhild Lothe, who with her supreme knowledge, enthusiasm and kindness has made it easy to thrive both scientifically and socially. Especially big thanks to post doc Guro Lind for being invaluable, both by helping out in the lab as well as with preparing manuscripts.

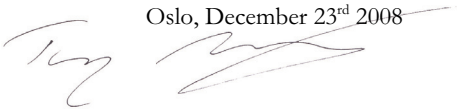
A big "grazie" to Professor Bruno Dallapiccola for inviting me to his excellent research facilities in Rome to work, and also to Irene Bottillo who with her excellent laboratory skills and including, friendly nature made my stay scientifically as well as socially very rewarding.

A big thanks goes to co-authors and collaborators for very good biological material, clinical data and scientific input. I would also like to acknowledge Professor Arild Nesbakken for being my official link to the Faculty of Medicine, as well as contributing with valuable feedback on the clinical aspects of the thesis.

Kudos to all my good friends and colleagues in the department who have meant the world to me. Special thanks to Helge and Marianne for making my day as we've been my office mates for 5 years. An important factor in completing this thesis is the personnel the Narvesen kiosk who ensured that my blood-sugar was at a constantly elevated level. Jan-Ivar Hansen, my biology teacher at Bjørknes Privatskole was so in love with biology that it was impossible not to be the same. Thanks!

This thesis had never even been started if it wasn't for my wonderful wife, Hege and my parents. Her for making me not want to study medicine in Hungary, and them for teaching me to have high standards and big aims. I love you!

Oslo, December 23rd 2008



PREFACE

"Is it just me? Or does everyone come across genetics, genes, and DNA almost everywhere?"

Madeleine Albright

Cancer as a disease has probably been around since the dawn of multicellular organisms. The oldest written description of cancer is as early as the 17th century BC in Egypt[1], but traces of cancer has been detected as far back as in dinosaur fossils. According to the humoral theory, it was believed that it was excess of black bile that was responsible for cancer as a whole. If this theory is correct the production of this body fluid must either have increased rapidly over the last centuries, or increase with age, as one out of three people are diagnosed with cancer today. Now we know that the black bile-theory did not stand the tooth of time, and that cancer is caused by both genes and environmental factors which via genetic and epigenetic abnormalities make normal regulation of *e.g.* cell cycle, apoptosis and signaling go awry.

With improved life expectancy and a more sedentary lifestyle comes increased lifetime risk of being diagnosed with cancer. Only in very rare occasions the primary tumor itself is fatal, but when the tumor metastasizes to distant and vital organs the mortality increases rapidly. As it takes time from the first alterations via precursor lesions and the primary - to a fully metastasized cancer, it leaves a window of opportunity in which the evolving cancer can be detected and the person cured if the suitable tool or biomarkers is at hand.

In September 2007, 4 years after the completion of the human genome project, the first individual genome was published[2]. Just a few moths later, the genome of James D. Watson, the co-discoverer of the double helix, was published using modern ultra high-throughput

technology[3]. This new technology, which dramatically reduces cost and time consume, paves the way for individualized medicine and treatment. The price for sequencing an individual genome is at the time being approximately $60.000. The ultimate goal is the $1000 genome which will make it possible for the average Joe to be deciphered. By then it will surely be easier to predict prognosis, treatment response and disease susceptibility among patients than what is the case today. Our job until then is to provide the community with good and trustworthy data on which future diagnoses and treatments can be determined.

LIST OF PAPERS

Ia Guro E Lind, **Terje Ahlquist**, Ragnhild A Lothe. DNA Hypermethylation of MAL: A Promising Diagnostic Biomarker for Colorectal Tumors. Gastroenterology. 2007. Apr;132(4):1631-2

Ib Guro E Lind, **Terje Ahlquist**, Matthias Kolberg, Marianne Berg, Mette Eknæs, Miguel A Alonso, Anne Kallioniemi, Gunn I Meling, Rolf I Skotheim, Torleiv O Rognum, Espen Thiis-Evensen and Ragnhild A Lothe. Hypermethylated MAL gene – a silent marker of early colon tumorigenesis. 2008. Journal of Translational Medicine. Mar 17;6:13.

II **Terje Ahlquist**, Guro E Lind, Vera L Costa, Gunn I Meling, Morten Vatn, Geir S Hoff, Torleiv O Rognum, Rolf I Skotheim, Espen Thiis-Evensen and Ragnhild A Lothe. Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers. 2008. Molecular Cancer – *In press*

III **Terje Ahlquist**, Irene Bottillo, Stine A Danielsen, Gunn I Meling, Torleiv O Rognum, Guro E Lind, Bruno Dallapiccola and Ragnhild A Lothe. RAS signaling in colorectal carcinomas through alterations of RAS, RAF, NF1 and/or RASSF1A. 2008. Neoplasia. Jul;10(7):680-6.

IV **Terje Ahlquist**, Ellen C Røyrvik, Marianne Merok, Gunn I Meling, Annika Lindblom, Xiao-Feng Sun**,** Georgia Bardi, Arild Nesbakken and Ragnhild A Lothe. Identification of *RCC2* as a prognostic marker among multiple gene mutations in colorectal cancer with defect mismatch repair. *Manuscript.*

Associated paper:

Ellen C Røyrvik, **Terje Ahlquist**, Torbjørn Rognes and Ragnhild A Lothe. Slip slidin' away: a duodecennial review of targeted genes in mismatch repair deficient colorectal cancer. Critical Reviews in Oncogenesis. 2007 Dec;13(3):229-57.

SUMMARY

The present study includes four papers presenting novel genetic and epigenetic changes that contribute to the development of colorectal cancer. Furthermore, some of these changes are suggested as suitable biomarkers for early detection and prognosis.

Through a novel experimental strategy *MAL* was identified as a potential epigenetically deregulated gene in colorectal cancer. Here *MAL* is identified as a target of tumor-specific methylation defined by high methylation frequencies in most adenomas (71%) and carcinomas (80%), and rare methylation in normal mucosa (4%). We show that positive methylation score depends on the assay design, that promoter hypermethylation close to the transcription start site is associated with reduced mRNA-expression of the *MAL* gene, and that the protein is seemingly absent in all tumors. We have therefore, most likely, identified a highly suitable diagnostic biomarker.

By comparing the methylation status of 11 selected genes in normal mucosa samples, precursor lesions and carcinomas from the large intestine we show a steady increase in "total" methylation in parallel with the tumorigenesis. A clear association is seen between frequent hypermethylation and microsatellite instability (MSI) of carcinomas, and indeed five genes (*CRABP1, MLH1*, *NR3C1*, *RUNX3,* and *SCGB3A1*) are shown to be MSI identifiers.

The mitogen activated protein kinase pathway is one of the most frequently affected signaling pathways in human cancer. Mutations of the oncogenes *KRAS* and *BRAF* are especially common in colorectal cancer. A negative regulator of KRAS is neurofibromin, encoded by the neurofibromatosis type 1 gene, *NF1*. The entire coding sequence of 61 exons is for the first time analyzed for mutations in a series of colorectal tumors. We found that *NF1* is primarily mutated in the non-coding introns, close to the intron-exon boundaries, and was typically found in MSI tumors. The *NF1* mutations identified are in contrast to known mutation spectra of NF1 patients and of their tumors. Overall, changes were detected in one or more of *KRAS*, *BRAF*, *RASSF1A* and *NF1* in >70% of all the analyzed carcinomas which further underline the importance of this pathway in cancer.

The subgroup of carcinomas with the MSI phenotype is known to have a better prognosis than a consecutive unselected patient series. We analyzed 41 genes for mutations in two clinically independent series of MSI-tumors and evaluated their prognostic information. We identified mutation status of *RCC2* as an independent prognostic marker, which discriminate between good and poor prognosis among patients with a localized disease.

GENERAL INTRODUCTION

Genetic and epigenetic inheritance

"Genetic inheritance is the biological process whereby genetic factors are transmitted from one generation to the next"

The principles behind genetic inheritance was elegantly described by Gregor Mendel in the mid 19th century, where he showed that an offspring inherit one allele of each gene from each of its parents, and that the encoded information from these make out the phenotype[4]. Although it took 34 years until his work was re-discovered and appreciated, the Mendelian heritage has ever since been widely accepted as the default way of genetic inheritance. The fact that DNA is the crucial compound of this inheritance was not shown until 1952, almost a century later[5], a year before the DNA structure itself was resolved[6]. Since then, the knowledge of inheritance has exploded. Some features are now known to be inherited in a non-mendelian manner, *i.e.* not according to Mendel's laws of inheritance. We also know that it is possible that information which is not coded by the DNA sequence itself can be passed to the next generation.

This exception from the Hershey and Chase findings in 1952 is called epigenetics. There are two kinds of epigenetic inheritance; cellular/somatic, the one from a cell to its daughter cells; and transgenerational/germline, from one parent to its offspring, the latter a recent and debated finding[7]. When epigenetic information is passed from a mother cell to its daughter cells, methylation marks are kept on the initial DNA strand during replication and cell division, causing the epigenetic pattern to be re-established in the daughter cell[8].

"Epigenetic inheritance is the transmission of information from a cell or multicellular organism to its descendants without that information being encoded in the nucleotide sequence of the gene."

Cancer as a genetic and epigenetic disease

There are innumerable components in the human cells involved in keeping close control on homeostasis, and a very simplistic view on cancer is that it is a disease caused by a skewed ratio in proliferation and apoptosis, favoring cell proliferation[9]. It does not need to be a huge growth-advantage in order to abolish homeostasis and cause transformed cells to take over the population. An 1% increase in the proliferation-apoptosis ratio will cause the affected cell to go from 0.001% to 99.9% of the population in 5 years given that the cell divides every 24 hours[10]. This illustrates the theory of clonal expansion which say that increased fitness of a cell will lead to clonal selection[11]. The growth advantage can occur by various mechanisms, and six hallmarks which cancer cells must acquire in order to reach malignancy has been suggested, including resistance to apoptosis, self-sufficiency in growth signals, insensitivity to anti-growth signals, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis[12] (Figure 1).

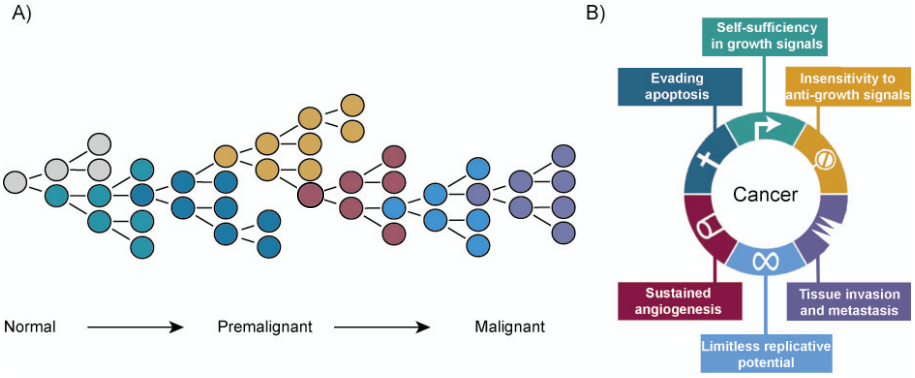


Figure 1. Biology of cancer. (A) Most cancers are believed to undergo a clonal expansion. The genetic or epigenetic events may be of any kind giving the cell a selective advantage. (B) Ultimately, these changes may fulfill Hanahan and Weinberg's six hallmarks of cancer. Courtesy of RI Skotheim.

The genes involved in maintaining homeostatis has traditionally been divided into two categories, proto-oncogenes and tumor suppressor genes. Mutations or chromosomal rearrangements may either activate oncogenes, causing them to drive *e.g.* cell proliferation even in absence of proliferative signals, and/or inactivate tumor-suppressor genes, leading to destruction of important checkpoints at which the cell may have time to perform DNA repair, align chromosomes or just wait for the beneficial environmental conditions for cell proliferation. Additionally, maintenance genes, or stability genes, are important for homeostasis. Under normal circumstances these genes keep the genetic errors in the cell to a minimum[13]. The fact that 30% of our genes encode proteins involved in regulating the DNA fidelity clearly shows the importance of these stability genes[14].

The impact of epigenetics on tumorigenesis is now widely demonstrated, and it is reckoned to play just as an important role in tumor development as genetics[15]. Promoter hypermethylation is a well known mechanism which may cause reduced or absence of expression of tumor suppressor and DNA repair genes while hypomethylation may activate a proto-oncogene into an oncogene[16]. As of this, both genome wide hypomethylation, and gene specific hypermethylation are involved in colorectal tumorigenesis[17].

DNA mutations in cancer

Mutations include all stable, irreversible changes that alter the primary DNA sequence, and are seen in less than 1% of a population. Today the term "mutation" is often thought of as subtle sequence alterations such as base substitutions and small insertions and deletions (indels), but in genetics, gross chromosomal alterations such as chromosomal translocations, gene amplifications, large gene deletions, and gains and losses of large chromosomal stretches were also included in the term, and that the gross and subtle changes were two different classes of mutations[18]. The different types of mutations and their timing in the cell cycle are illustrated in Figure 2. There are different categories of subtle mutations: point mutation, defined by the exchange of a single base, includes both missense mutations (exchange of a base giving an amino acid change) which can affect protein folding and function, and non-sense mutations (exchange of a base which inserts a premature termination codon) leading to a truncated protein. If a point mutation does not give an

amino acid change it is referred to as a silent mutation. Silent mutations can still have an impact on the correct transcription of genes if they occur close to an intron-exon boundary. If so, it may affect correct splicing of the exons and be referred to as a splicing mutation. The first two bases up- and downstream of an exon are called the acceptor and donor site and consist of GT and AT, respectively. Alterations in any of these four bases will lead to exon skipping and a shorter protein[19]. Indels lead to frameshift mutations (except if the indel consists of a number of bases dividable by 3), which in most cases leads to a premature termination codon and a truncated protein.

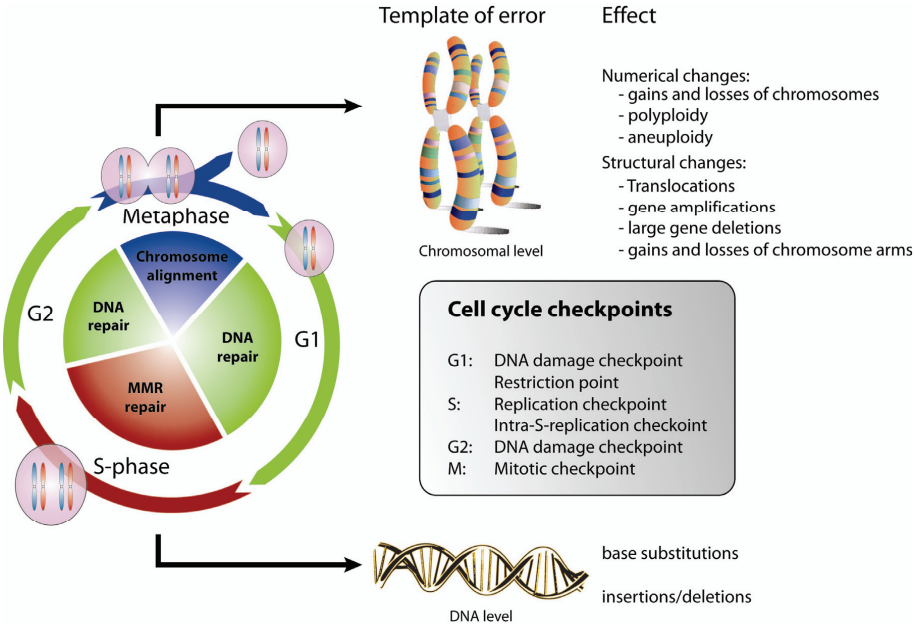


Figure 2. Cell cycle and timing of mutations. The different kinds of mutations occur at different stages in the cell cycle. During G1, chemical compounds and UV-irradiation can cause DNA damage which needs to be detected and repaired before the re-synthesis if mutations are to be prevented. Just before the G1/S-transision, the cell enters the restriction point, in which passing irreversibly commits the cell to undergo DNA re-synthesis and cell division. Only optimal conditions and minor DNA errors not detected by the repair machinery makes the cell pass this point. If not, it will enter a quiescent G_0-state (not shown). During S-phase several DNA-repair systems work in order to obtain high DNA fidelity. Further, the same DNA damaging agents who cause damage in G1 are present in G2, which means that before cell division, the cell monitors the DNA quality and the chromosome alignment before dividing.

All bases have the same theoretical chance of being affected by mutations. Still, in human cancers we see a clear accumulation of mutation of specific bases in specific genes,

exemplified by frequent mutations in specific regions of *TP53*[*], *KRAS*, *BRAF* and *PIK3CA*[†]. This can be explained by selection pressure and clonal growth advantage[11]. If a mutation gives the cell a growth advantage it will have higher fitness compared to surrounding cells, leading to selection in accordance with the evolution theory[20]. On the same basis, , a cell with a "non-important" mutation, often denoted passenger mutation, will not have increased likelihood of passing the acquired trait to its daughter cells. In the largest available mutation database, Catalogue Of Somatic Mutations In Cancer (COSMIC), ~4800 of our 20-25000 genes are included[21], a number which also includes a large number of genes with no observed mutations. The number of genes which is essential for tumor formation is much lower, and it has been estimated that just more than 1% of the human genes can be attributed "cancer critical" genes, as these are known to be involved in carcinogenesis[22].

DNA methylation

Methylation is one of many possible chemical modifications of the double helix. At a CpG dinucleotide, a cytosine followed by a guanine, a methyl group can covalently bind to the 5'position of cytosine, constituting a 5-methylcytosine, first described in 1948[23]. This will serve as a recognition mark for methyl binding proteins. The CpG dinucleotide is greatly underrepresented in the human genome due to spontaneous deamination of 5-methylcytosine throughout the evolution, causing a cytosine to thymine change that escapes the DNA repair systems[24]. Hence, the remaining CpG-sites are mainly located in areas with a close to theoretically expected amount of CpG, called CpG-islands, but are also present within repetitive sequences. These islands are located in the 5'-ends of about half of our genes and are under normal circumstances protected against methylation[25].

Both establishment and maintenance of methylation marks are governed by DNA methyl transferases (DNMTs). During DNA replication hemi-methylated DNA is recognized by DNMT1, and the methylation marks of the initial strand are copied to the nascent DNA strand, ensuring a faithful inheritance of DNA methylation patterns in the daughter cells. This process is called maintenance methylation. DNMT3A and DNMT3B handle the *de novo*

[*] IARC TP53 database – http://www-p53.iarc.fr/
[†] Catalogue of Somatic Mutations in Cancer database – http://www.sanger.ac.uk/genetics/CGP/cosmic/

methylation during embryogenesis, and by such establish novel methylation marks on the DNA[8;26]. Some genes, including imprinted genes, are inherited via the germ line[27;28]. How epigenetic information can be passed via germ cells is still unclear as a substantial epigenetic re-programming takes place during early embryogenesis, often thought of as an erasure of methylation marks.

Promoter hypermethylation impair gene expression by preventing transcription factors to bind to DNA. This occurs in two ways, either by direct inhibition of the binding of transcription factors to the methylated sequence[29;30], or via recruitment of proteins with a methyl binding domain (MBD protein family) which leads to condensation of the chromatin structure by the means of histone deacetylation (reviewed in [31])(Figure 3). Knocking out MBD-proteins with RNA interference re-introduces transcription of initially inactive genes without altering the methylation status, thereby clearly demonstrating the role of MBDs as a link between promoter hypermethylation and gene expression[32]. Still, the role of mutations in either of the MBD genes are unclear[33;34].

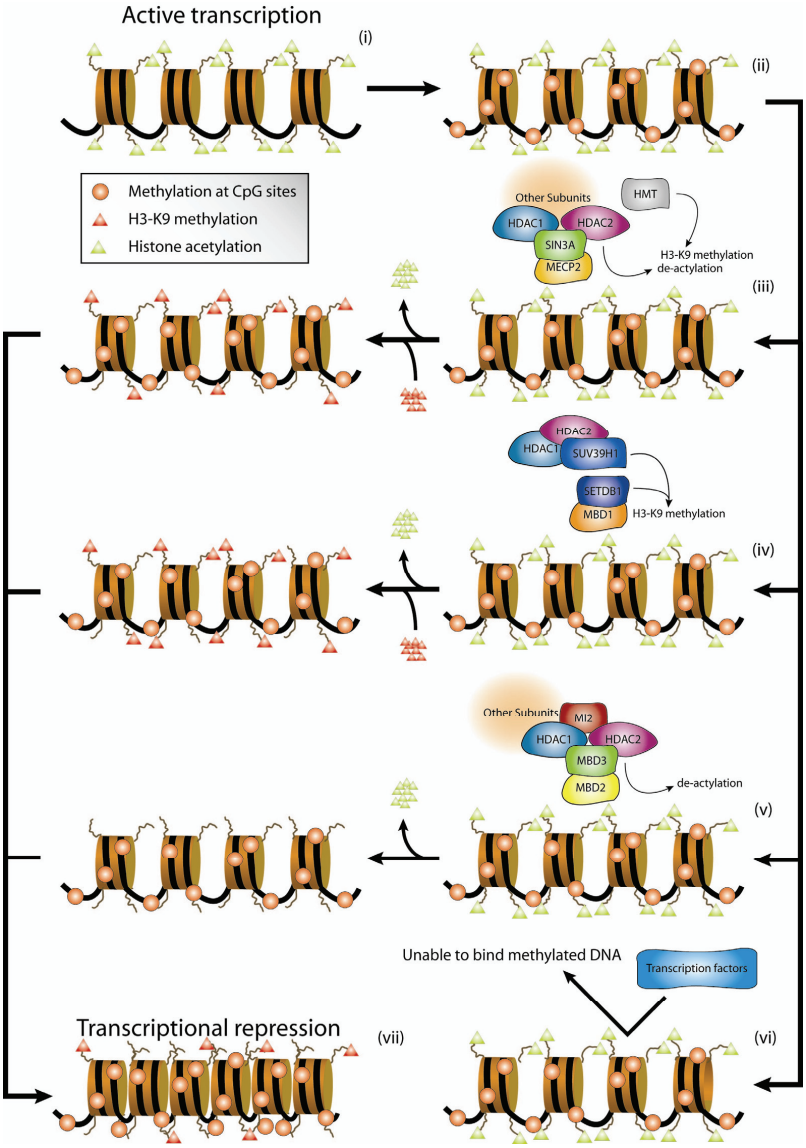


Figure 3. Methylation effect on chromatin. In an expressed state, the chromatin structure in proximity to the gene is loosely packed, a state recognized by acetylation of the histone tails (i). The adding of methyl groups to the CpG sites (ii) is associated with a more densely packed chromatin and can be obtained in several ways. A family of proteins with a methyl-binding domain (MBD) binds to the DNA-bound methyl groups. Depending on which protein that binds, the binding partners and the effects vary slightly (iii – v). The net effect is that RNA transcription is impaired due to blocking of transcription factors, either by inhibiting the direct binding (vi) or by condensation of the chromatin structure (vii).

DNA is most of the time wrapped around an octamer of histones, which together make up the nucleosome (see Figure 4 for DNA packing strategies). The octamer is comprised of two of each of the histones H2A, H2B, H3 and H4. Methylation of DNA is accompanied by post-translational modifications of histones which modulate DNA function, regulate chromatin structure and determine the transcriptional state of the DNA wrapped around it[31]. The sum of post-translational modifications of amino-terminal tails of histones constitutes the histone code[35], and include acetylation, methylation, phosphorylation and ubiquitinylation[36]. Certain modifications such as methylation of Lys4 of H3 (H3-K4) is associated with active gene expression, while others, like methylation of Lys9 of H3 (H3-K9) is associated with transcriptional silencing[8]. There is a whole range of different modifications, and much effort is aiming at fully deciphering the histone code.

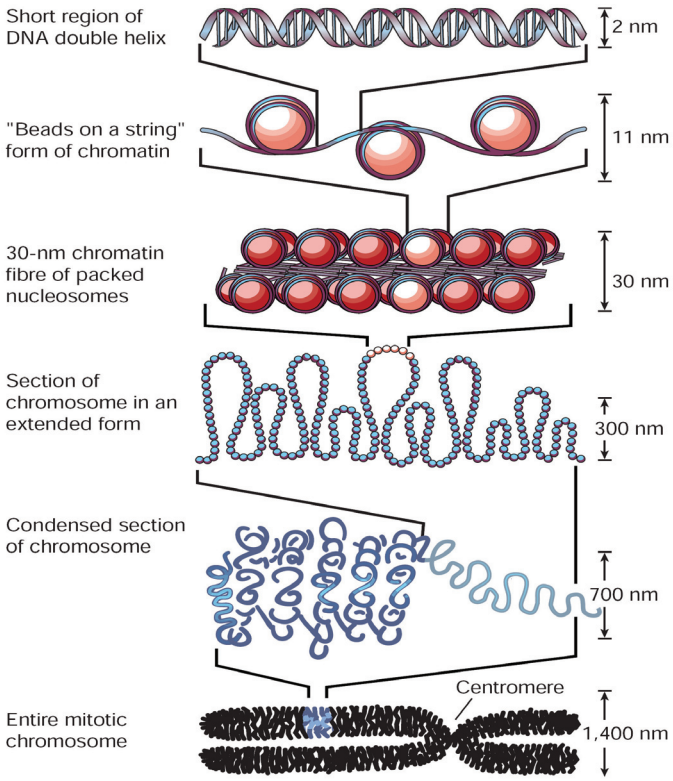


Figure 4. Packaging of DNA. 147 bases of the double helix is wrapped twice around the octamer of histones (nucleosome), with a short stretch of linker DNA connecting the nucleosomes, resembling beads on a string. The histone tails are available for post-translational modifications at this step, and depending on the modification the DNA can be more or less accessible for transcription factors. The nucleosomes are further packed into fibres of nucleosomes which is further condensed until the ultimate condensation, the visible chromosome during mitosis. Figure taken from [36].

Even though this study focuses on the role of DNA methylation in cancer, its role in normal life and development must also be addressed. During mammalian embryogenesis both the paternal and the maternal DNA in the zygote undergo extensive erasure and reprogramming of DNA methylation[8], resetting close to all methylation marks. The reason for this re-methylation process is unclear, but it has been suggested that it will lead to decondensation of chromatin and activation of transcription of genes which are important in early development[8].

Both X-chromosome inactivation and genomic imprinting are closely regulated by methylation. In contrast to most of the human genes which are expressed in a diploid manner, imprinted genes and the X-chromosome are only present in one parental copy, while the other is epigenetically silenced. The majority of imprinted genes identified so far are involved in growth, and imprinting might be a strategy to balance the maternal and paternal demands on the rate of fetal growth[37;38]. *IGF2* is an example of an imprinted gene, and while hypermethylation of gene promoters are associated with lost gene expression, studies have shown that hypermethylation of repressor elements within and upstream of the insulin-like growth factor 2 (Igf2) gene in mice increases the expression as proteins involved in the Igf2 repression are now unable to bind to the sequence.[39;40].

DNA methylation is also involved in silencing of repetitive and viral sequences present in our genome[41]. At least 35% of our genome is constituted by tranposons, viral DNA and other parasitic sequences[42], and the human cell protects itself from this by methylation-induced inactivation. While most of the CpG islands remain unmethylated under normal circumstances, the majority CpG-sites outside the CpG islands are methylated[25], and much of this methylation can be explained by this protective silencing strategy. Figure 5 summarizes the difference in methylation features between a normal and a cancer setting.

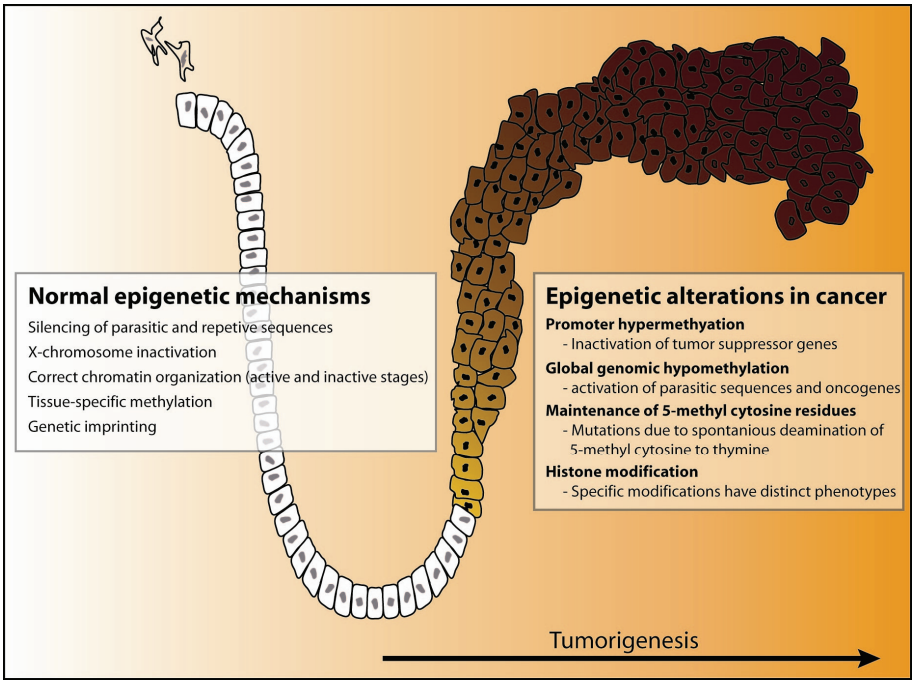


Figure 5. DNA methylation effects in normal and cancer cells. Methylation has important functions in normal development and normal cells as shown to the left. The effects of the loss of epigenetic control in cancer are shown in right side of the figure.

DNA methylation and cancer

Ever since the discovery of hypomethylation in human tumors in 1983[43], it has been clear that epigenetics, and in particular methylation, plays an important role in cancer development. Even though hypomethylation triggered the interest of epigenetics and cancer, it is promoter hypermethylation that have stolen most of the headlines thereafter, and especially subsequent to the finding of promoter hypermethylation of the *RB* gene, the symbol of Knudson's two-hit hypothesis from twenty years prior[44], when the link between methylation and reduced transcription was found[45]. In spite of the reserved suggestion of promoter hypermethylation being an infrequent event in human cancers, the article paved way for an even stronger focus on methylation as an important factor in tumorigenesis[45;46].

A rapidly growing field is the epigenetic control of micro-RNA (miRNA). miRNAs are endogenous non-coding small RNAs which regulate gene expression in a sequence-specific manner, and can function as either oncogenes or tumor suppressor genes[47]. Compelling evidence now show that miRNAs are subject to both hypo- and hypermethylation in a tumor and tissue-specific manner[48]. New studies also suggest that hypermethylation can mimic small chromosomal deletions or loss of heterozygosity by means of long range epigenetic silencing (LRES)[49;50]. The concept of LRES is no longer "one methylated CpG island – one silent gene", but rather involves large regions which may include several genes[49]. In breast cancer, the HOXA cluster has been shown to be subject to LRES, in which approximately 100kb DNA undergoes epigenetic inactivation[51].

Colorectal cancer

Incidence, treatment and outcome – Nature versus Nurture

Each year over 1 million people are diagnosed with colorectal cancer (CRC) worldwide[52], including 3500 new Norwegian cases. It is an age-related disease, in the sense that it primarily affects older individuals (Figure 6). CRC is the most common form of sex-independent cancer[‡], and the incidence is increasing (Figure 6). The disease is more prevalent in the industrialized countries (Figure 7)[§]. Understanding both the initial and progressive steps of the disease is important in order to learn how and who to treat as knowledge of early alterations will provide information on how to detect the disease at an early, curable stage, while the later changes may be exploited to give information on treatment response. The substantial differences in CRC incidence between the industrialized and developing countries suggest that the disease is strongly affected by life-style and environmental factors. Several risk factors have been suggested and include obesity, high-fat diet, tobacco, alcohol and lack of physical activity[53]. Other factors have a preventive effect on CRC, *e.g.* exercise, high intake of fiber and vegetables as well as sufficient amounts of different nutrients. However, the literature is inconsistent on the matter, and the different mechanisms of risk factor exposure are still unknown[54].

‡ The Norwegian Cancer Registry – http://www.kreftregisteret.no
§ International Agency for Research on Cancer web pages - Cancer Mondial – http://www-dep.iarc.fr

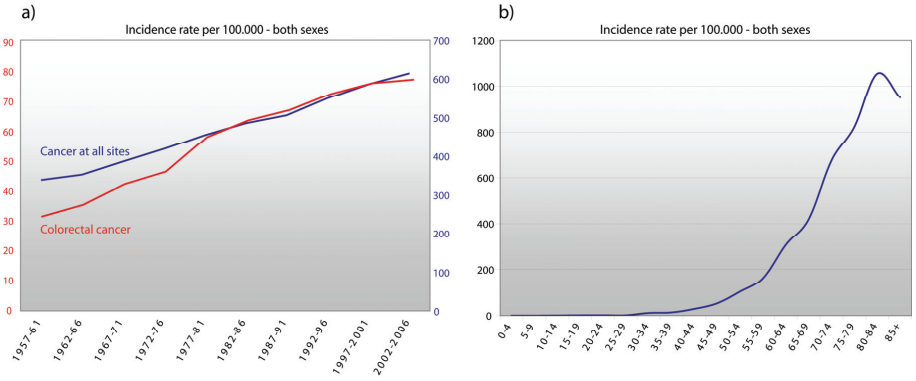


Figure 6. Incidence rate for colorectal cancer compared with other types of cancer and age.
a) Age-adjusted incidence rate for CRC compared to cancer in general in a Norwegian population. b) The incidence of CRC is plotted against age at diagnosis. Raw data was obtained from the Norwegian Cancer Registry – Cancer in Norway 2006.

The majority (70%) of CRC occurs sporadically, and only ~5% of all registered CRC are due to single hereditary components, such as Lynch syndrome (OMIM #120435), formerly known as hereditary non-polyposis colorectal cancer (HNPCC), and familial adenomatous polyposis (FAP, OMIM #175100). The remaining ~25% of CRC account for families with an accumulation of cancer, but with no known clear-cut Mendelian inheritance[55], and may be multifactorial. Some studies claim that the fraction of heritable CRCs are as much as 35%[56].

Both Lynch syndrome and FAP are autosomal dominant disorders with a relative life-time risk of developing CRC of 80% and 100%, respectively[57]. Lynch syndrome is caused by a germ line mutation in one of the components of the DNA mismatch repair system (MMR), most commonly *MSH2* or *MLH1*[14] (see page 28-30 for MMR details). FAP is caused by germ line mutation in the *APC* gene, a central cytoplasmic complex involved in degradation of β-catenin (*CTNNB1*) in the WNT signaling pathway.

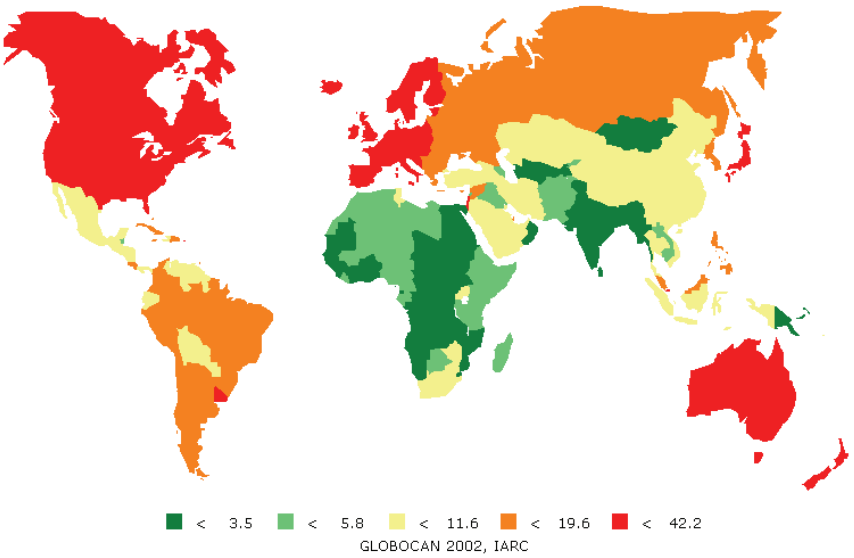


Figure 7. Colorectal cancer age-standardized incidence rate per 100.000 for women. The map for males shows an identical trend; that the industrialized part of the world has a high incidence compared to the developing countries.

Survival among CRC patients depend heavily on tumor stage at time of diagnosis (Figure 8). Localized disease is associated with a relatively good prognosis (88% five-year relative survival) while patients with distant metastasis have a five-year relative survival of only 8%. Across all stages the five-year relative survival is approximately 57%. Another factor which predicts patient survival is genetic instability in the tumor, which traditionally is divided into chromosomal instability (CIN) and microsatellite instability (MSI), and patients with a MSI tumor is associated with a better - while those with CIN tumors have a worsened survival[58-60]. A third more recently described instability, called CpG island methylator phenotype (CIMP), includes a subgroup of CRC with epigenetic instability which drives tumorigenesis in a similar way as CIN and MSI[61]. A detailed description of these phenotypes is presented on page 28-32.

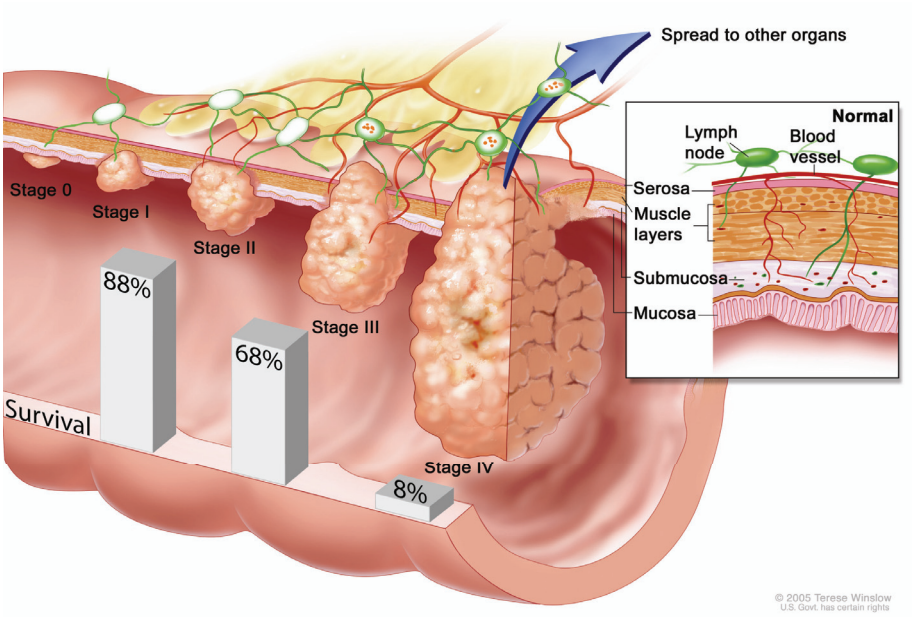


Figure 8. Tumor staging and survival. Colorectal tumors are staged based on degree of localization. Stage 0 is often referred to as carcinoma *in situ* and has not penetrated the mucosa boundaries. Stage I is considered localized as the tumors has not penetrated the serosa. Stage II tumors are also considered local, but have penetrated serosa and may have grown into pericolic fat tissue. Patients with localized tumors have a 5-year survival of 88%. Stage III tumors are considered regional as local lymph nodes are infiltrated. Five-year survival among such patients is 68%. Patients with a stage IV tumor have metastasis to distant organs such as the liver, and have the worst prognosis as only 8% survive. Survival data has been added to this illustration by National Cancer Institute.

Genetics and epigenetics of CRC – a timeline

Colorectal cancer is one of the most extensively studied cancer types over the last century resulting in small and bigger leaps of knowledge. The timeline in Figure 9 presents some of the most important discoveries for this disease as well as for cancer in general.

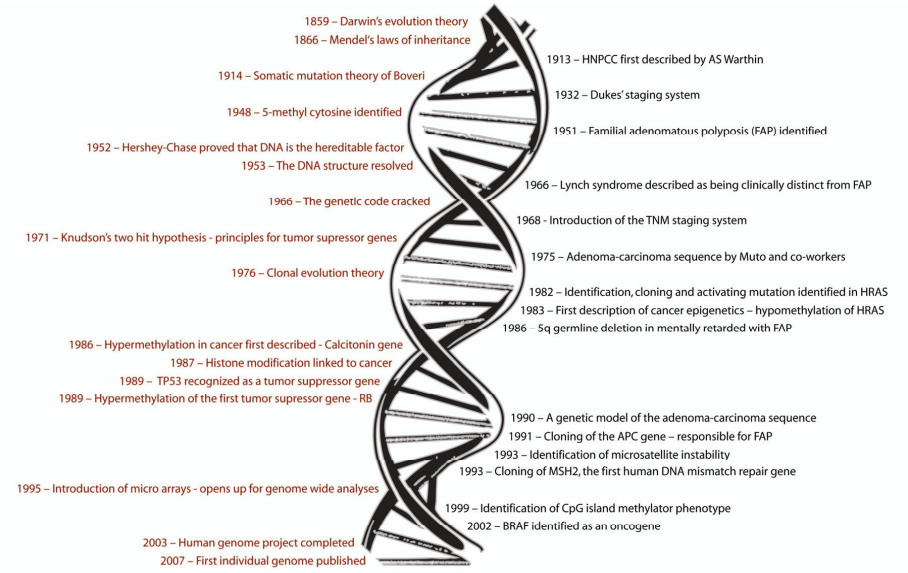


Figure 9. A timeline of (colorectal cancer) genetics. The left hand side of the helix (in red) lists general genetic hallmarks while the right hand side (in black) includes hallmarks in colorectal cancer research.

Molecular and morphological developmental pathways in colorectal cancer

The adenoma-carcinoma model was first presented by Muto and co-workers in 1975[62]. It described how colorectal cancer developed through different histopathological steps, from the first abnormality, via increasingly dysplastic adenomas to carcinomas, and has ever since been a widely accepted paradigm for colorectal carcinogenesis. In 1990, Fearon and Vogelstein added genetic information to this model and suggested that alterations in at least four of the presented genes were sufficient for developing cancer[63]. It has later been shown that the order of events is just as important for the cancer development as the number of alterations[64;65].

According to the adenoma-carcinoma model, adenomas can develop into either microsatellite- or chromosome instable tumors depending on the genetic makeup of the tumor[64]. Mutation in the tumor suppressor gene *APC* is considered to be the initiating event, or "gatekeeper mutation" in colorectal tumorigenesis, and has been reported

inasmuch as 80% of colorectal tumors [64]. Activation of the proto-oncogene *KRAS* is also considered an early event, and is present in 35% of colorectal carcinomas[**]. *TP53*, called "the guardian of the genome"[66] is a critical component in sensing DNA damage and stress signals, and is the most frequently altered tumor suppressor gene in human cancer in general. In colorectal cancer it is not altered until late in the development, in advanced carcinomas, and close to 50% of all CRC have mutations in *TP53* at this stage[††]. According to this model, microsatellite instability is a late event as it is only seen in advanced adenomas and subsequent stages[64].

Compelling evidence in recent years have questioned this paradigm and turned the scale towards adenomas giving rise only to chromosome instable and CIMP-negative tumors, while the previously "ignored" hyperplastic polyps (HPs) have re-established their malignant potential, as a subgroup of these, the sessile serrated polyps, are likely to give rise to MSI and CIMP tumors[67-70]. An equivalent to the adenoma-carcinoma sequence has now been suggested and includes the development from normal mucosa via HPs and serrated lesions to MSI-carcinomas. Typical for these cancers is that mutation in the proto-oncogene *BRAF* is considered the "gatekeeper", and that the dysregulation of the methylation machinery occurs before microsatellite instability (Figure 10)[71].

Hypermethylation in normal-appearing colorectal mucosa adjacent to the primary tumors has gained interest lately. This abnormality is thought to be *the* initiating event in tumorigenesis, as it presents a field in which the cells are especially susceptible for additional alterations. With a sufficing amount of alterations, the tumorigenesis is initiated. This phenomenon is called "field effect" or "field cancerization", and several genes have been suggested to generate such[72-75].

[**] Catalogue of Somatic Mutations in Cancer database – http://www.sanger.ac.uk/genetics/CGP/cosmic/
[††] IARC TP53 Mutation Database (Release 12) - http://www-p53.iarc.fr/

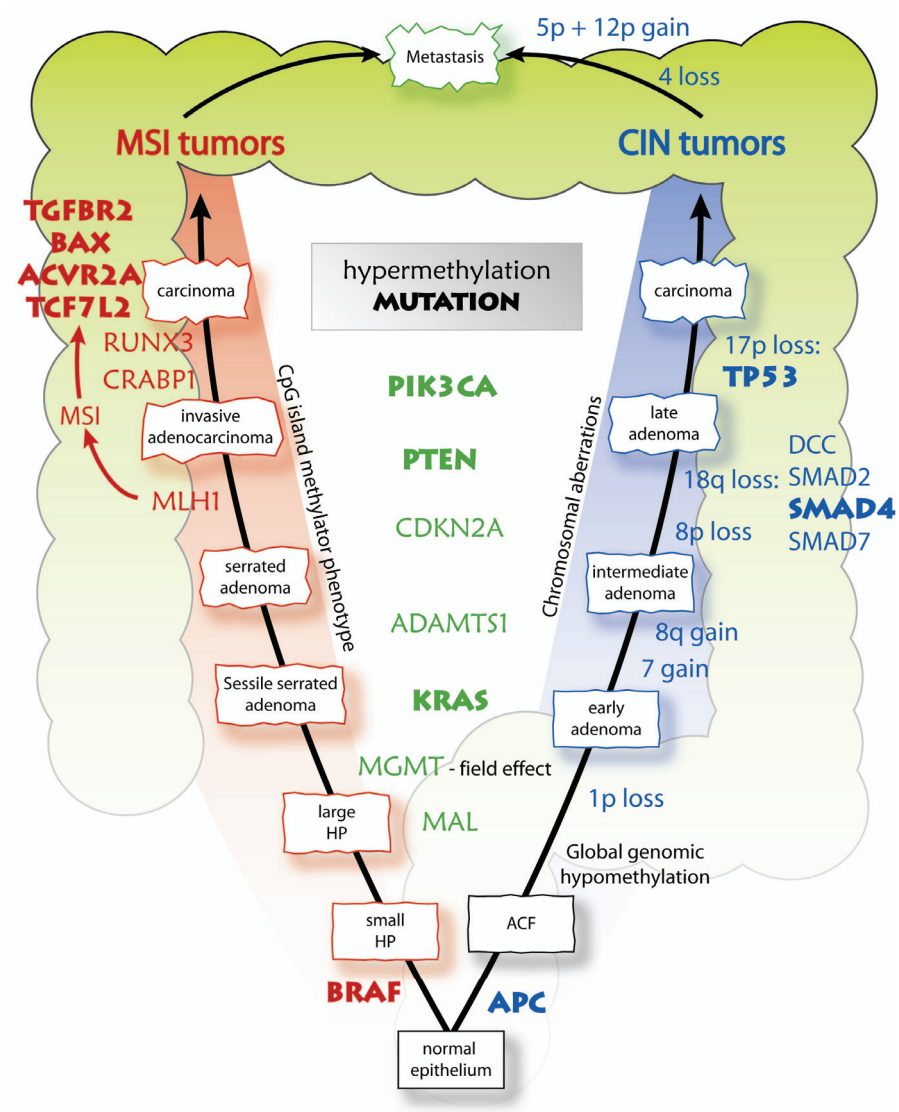


Figure 10. Molecular pathways to colorectal cancer. Colorectal cancer develops through several distinct histopathological steps, each of which is associated with different kind of alterations. It is now believed that MSI (in red) and CIN tumors (in blue) develop through two distinct pathways, the sessile serrated pathway, giving rise to MSI-tumors in the proximal colon, and the traditional adenoma-carcinoma pathway, giving rise to CIN tumors in the distal colon. While *APC* is considered to be the initiating event in the initiating event for CIN tumors, *BRAF* is one of the earliest recognized changes among the MSI precursors. The genetic complexity increases throughout CIN tumorigenesis due to the chromosomal instability, while CIMP leads to epigenetic instability among the MSI precursors, eventually affecting MLH1, which then causes MSI. With MSI, genes carrying repetitive units within their coding region are especially susceptible for mutations.

Instabilities involved in colorectal cancer

Chromosomal instability – CIN

Tumors with chromosomal instability are characterized by aneuploidy and numerous chromosomal aberrations. The majority of colorectal carcinomas (~80%) display CIN, and these tumors are most frequently found in the left, or distal, side of the colon[14]. CIN is associated with a worse prognosis in CRC with a hazard ratio of 1.45 as compared to CIN-negative tumors[60]. Also, mutations in *APC*, *KRAS* and *TP53* are more prevalent among CIN tumors[67].

Chromosomal instability should not be confused with chromosomal complexity, as a tumor with a complex karyotype not necessarily displays CIN. An unstable cell will not have the same karyotype after a few cell divisions, while a stable cell will do so, regardless of whether it has a normal or a complex karyotype[10].

The quest for the CIN-causing mechanism has been around for a long time. Genes involved in the mitotic spindle checkpoint, centrosome regulation, DNA damage- and replication checkpoints, cell cycle and telomere elongation has been mentioned as a probable cause for CIN, and it seems that proteins involved in regulating spindle-kinetochore interactions during mitosis have the strongest evidence to play a part[14]. In total, more than 100 candidate genes causing CIN has been suggested, but to date, the mechanisms behind CIN are still unknown[14].

Microsatellite instability – MSI

The concept of genomic instability has been known for a very long time[76], but MSI in human cancer was not identified until 1993[77-79]. Microsatellites are scattered around the human genome, and consist of repetitive units of 1-6 bases flanked by unique sequences[80]. MSI is defined as an event occurring when a germline microsatellite allele has gained or lost repeat units, leading to a somatic change in length[81]. The underlying cause of MSI is

defective DNA mismatch repair, which means that MSI is a marker for loss of DNA mismatch repair (MMR) activity.

Under normal conditions the MMR-system monitors DNA replication, where it recognizes and removes mispaired base-pairs. The gain or loss of repeat units is called insertion or deletion of bases, or indels in short. Indels are introduced when there is a transient, local dissociation of the daughter and parental DNA strands in a microsatellite and the two strands subsequently undergo re-annealing between misaligned repeat units[82;83], resulting in lengthening or shortening of the daughter strand. This phenomenon is called replication slippage[84].

The fidelity of DNA replication under normal conditions is 1 error per 10^9-10^{10} nucleotide synthesized[85]. The initial error-rate is much higher, but proof-reading enzymes and DNA repair systems reduces this a hundred-fold[86]. A malfunctioning MMR system will therefore bring the error rate of indels up to the initial level, giving rise to MSI.

Lynch syndrome is as mentioned caused by a germline mutation in one of the MMR-components (see page 22), usually *MSH2* or *MLH1*[14]. These individuals are characterized by an earlier onset of CRC compared with the sporadic cases, and the tumors are always of the MSI phenotype. In contrast, the sporadic microsatellite unstable colorectal tumors do very seldom carry a mutation in either of these genes. Instead, hypermethylation of *MLH1* is seen in 80% of the patients[87]. Unlike the CIN tumors, MSI tumors have a diploid or near-diploid karyotype[78]. They are primarily located in the right, or proximal, side of the colon, and are more prominent among women[88]. Initial studies claimed that MSI was associated with a better prognosis[58], which has been verified in a large meta-analysis[59].

It is already mentioned that microsatellites are scattered around the genome, but this is a truth with modifications. Microsatellites are *unequally* distributed in the genome, with the majority in non-coding areas. This is probably the result of selection against easily disrupted sequences in coding regions[80]. A general rule is that the shorter the repeat unit and the longer the repeat, the more unstable is the microsatellite[80;89]. More repeat units provides a higher number of correct bases which may stabilize the misaligned product[86]. In general,

there seems to be a fidelity threshold for homopolymer tracts, where those longer than 7 base pairs are considered much less stable than those of seven or fewer units[80].

With well over a thousand genes containing eight or more mononucleotide repeats in the genome[90] it is obvious that mutations in these genes will accumulate at a much higher rate among MSI tumors compared to tumors with a functional MMR-system. MSI is therefore often referred to as the mutator phenotype[91].

CpG island methylator phenotype – CIMP

While CIN tumors have a defect in the chromosomal segregation and MSI tumors have a defective MMR-system, the CIMP-tumors have a methylation machinery that has gone awry, causing epigenetic instability[92]. Tumors with CIMP are therefore associated with promoter hypermethylation of a large number of genes. It has previously been shown that 38% of CRC display neither CIN nor MSI, and it has been speculated whether this group has epigenetic events as the malignant driving force[93]. The CIMP phenotype overlaps in a large degree with MSI, resulting in many of the same clinical associations, such as proximal location and female gender of old age[92]. Overall, CIMP is associated with a poor prognosis compared to CIMP-negative tumors[94;95]. In addition, it has been shown to be an independent (from MSI and *TP53*^{mut}) positive predictor of survival benefit from 5-FU treatment[96]. On the genetic level, CIMP is strongly associated with the V600E mutation in *BRAF*, *KRAS*^{mut}, and *TP53*^{wt}[97;98], which has led to the suggestion of two CIMP-groups; CIMP1 which is associated with extensive hypermethylation, V600E mutation in *BRAF* and MSI; CIMP2, an intermediate group with less frequent hypermethylation than CIMP1 as well as *KRAS* mutation; and the CIMP negative group which contains very scarce methylation, microsatellite stable and wild-type *TP53* tumors[98](Figure 11). CIMP is inversely associated with mutations in *APC*, the hallmark of early events in adenomas[99], thus supporting the hypothesis that CIMP and MSI tumors originate from the sessile serrated polyps and not adenomas[100]. CIMP is not exclusive for CRC. It has also been shown in gastric-, lung-, liver- and ovarian cancer as well as leukemia[95].

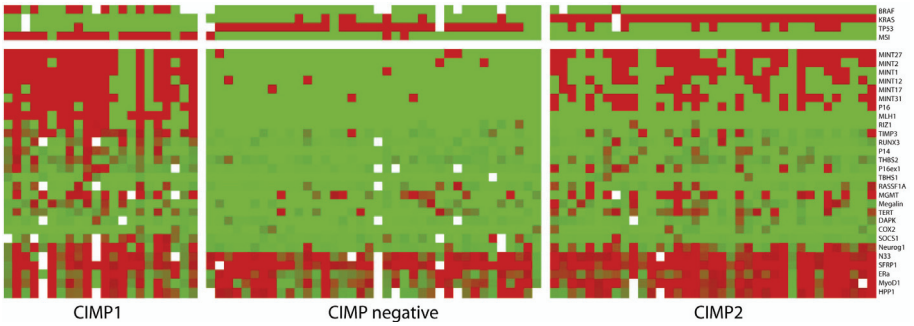


Figure 11. Integrated genetic and epigenetic analyses identify CIMP. With an integrative analysis of mutation status of BRAF, KRAS and TP53, MSI status and methylation status of a panel of genes, Shen and co-workers identified three distinct subgroups of CRC. CIMP1 carry *BRAF* mutation, MSI and wide-spread hypermethylation, CIMP2 carry *KRAS* mutation and intermediate degree of hypermethylation, while the CIMP negative group carry *TP53* mutations and scarce methylation (from [98]).

CIMP was first described in 1999[61] and became immediately an area of great dispute. Toyota and colleagues identified a group with concurrent methylation of several loci and another group with very little methylation. The group with frequent methylation was named CIMP. Several studies supported this new phenotype[97;101;102], while other studies failed to verify the bimodal distribution of methylated genes, thus turning down the CIMP-phenotype[103-105]. Much of the discrepancies seen between the different studies are likely due to both the panel of genes determining CIMP as well as the methodology[106;107]. Today, CIMP is accepted as a third phenotype, and the effort has turned to how to correctly classify the CIMP tumors[98;108;109], as well as determining the initial cause. CIMP markers will have to go through a similar process as the MSI marker panel (which initially was different from publication to publication) to reach a consensus marker panel [110].

Among the original objections against CIMP was that it overlapped extensively with the MSI phenotype, and that CIMP was plainly a subclass of MSI-tumors with extensive promoter hypermethylation[105]. In fact, according to the sessile serrated polyp-model described on page 26, it may prove to be the other way around, namely that CIMP precedes MSI, as CIMP may cause *MLH1* to be hypermethylated, which again leads to MSI.

The underlying cause of CIMP is still uncertain. Initially, the DNA methyl transferases (*DNMT1*, *2*, *3a* and *3b*) were considered good candidates, but to date, no mutations or

severe alterations are described in human cancer that can confirm this. On the contrary, germ line mutation in *DNMT3B* are known to cause ICF syndrome (Immunodeficiency-centromeric instability- facial anomalies syndrome, OMIM #242860), a disease associated with hypomethylation of small percentage of the genome[111], but without elevated cancer risk. Overexpression of *DNMT1* has been another suggestion for CIMP development, but different technologies give inconsistent results[95]. Also it has also been speculated whether loss of methylation boundaries may cause CIMP as methylation naturally increases with age. This age-specific methylation is under normal circumstances kept within its boundaries. In cancer, these boundaries can be lost, leading to spreading of methylation and so-called cancer-specific methylation[92][#].

In summary, MSI is the only of the three different instabilities in CRC of which we know the cause, although it seems that the real cause of sporadic MSI is CIMP and that MSI is a mere consequence of CIMP. (refer Figure 10).

Signaling pathways

The critical feature in tumorigenesis is often not the altered gene *per se,* but rather its involvement in a complex signaling network. Alteration in one gene may therefore trigger a cascade of signaling factors, such as when *KRAS* activate the mitogen activated protein (MAP) kinase pathway[112]. Hence, from a cancer cell perspective, it is more beneficial to alter a gene involved in signaling networks instead of a "lonesome wolf".

Some pathways play a more prominent role in tumorigenesis than others, possibly because it is easier for the cell to "highjack" them. The WNT pathway is an example of this. It controls many events during embryogenesis, including regulation of proliferation, morphology, motility and cell fate[113]. It is believed to be dysregulated inasmuch as 90% of colorectal cancers and *APC*, *AXIN2*, *CTNNB1* and *TCF7L2* are the most frequently altered components[114;115]. The MAP kinase pathway, involving *KRAS*, *BRAF*, *RASSF1A* and *EGFR* as common targets in cancer is another example and is deregulated in at least one

[#] Jingmin Shu – unpublished – presented on Cancer Epigenetics, AACR Special Conference, Boston, USA, May 2008

third of human cancers[112;116]. PI3K – AKT pathway is a key regulator of many important cellular activities, including proliferation, cell growth, survival, and metabolism. It is highly conserved from *C.elegans* to humans, indicating an essential function in the cell [117]. The pathway is overactive in 30% of human cancers and pathway members, including *AKT2*, *IRS2*, *PIK3CA*, *PTEN*, and *PDK1* have been found altered in different kinds of cancer (reviewed in ([118]). Transforming growth factor β (TGFβ) evolved to regulate epithelial and neural tissues, the immune system, and wound repair in vertebrates[119]. This two-faced pathway is of special interest in a cancer perspective as it works in both an anti- and a pro-tumorigenic fashion. The anti-tumorigenic effects of TGFβ is exerted via the downstream signaling targets which includes important cell-cycle checkpoint genes such as *CDKN1A* (p21), *CDKN1B* (p27) and *CDKN2B* (p15)[120], genes with tumor suppressive activity. The pro-tumorigenic function is due to TGFβ being a potent inducer of epithelial-mesenchymal transition (EMT). Cells undergoing EMT acquire motility and invasive properties, traits essential for metastasizing cancers[121]. This leaves two possibilities. One is to inactivated the whole pathway, and in that way get rid of the tumor suppressing activity. The other is to specifically knocking out just the tumor suppressing arm of TGF-signaling: As with the first possibility this will eliminate the suppressing activity, but in addition the cell may exploit the pro-tumorigenic phenotype of the pathway to enhance tumor growth and invasion[119]. It is therefore of no surprise that TGFβ signaling is among the most commonly altered signaling pathways in human cancers[120].

Clinical challenges for CRC

Early detection and screening programs

There are several clinical challenges with colorectal cancer. The disease is equally frequent among men and women and that the overall patient survival is poor. The fact that survival depends heavily on stage at time of diagnose indicate that great effort should be made in order to diagnose patients as early as possible in the tumor time-line. Different approaches are suggested in order to improve early detection. Population-wide colonoscopy or sigmoidoscopy has been shown to reduce the relative risk of developing CRC by 80%[122]. Even though colonoscopy is considered the golden standard in colon cancer screening with estimated benefits as high as 90% in reducing mortality, only a few countries have a

population-wide screening program. The obstacles of this method are the cost and availability of trained personnel, as well as reduced compliance over time. Randomized trails designed to determine the efficacy with regards to incidence and mortality are therefore needed in order to justify the use of colonoscopy as a population based screening program[123].

The presence of occult blood in feces is considered an indication of CRC, and a fecal occult blood test (FOBT) is already in use as a screening program in several countries. FOBT is shown to reduce mortality by CRC with 15-33%, but the need for biannual testing has resulted in a decrease in patient compliance over time and lower the effect of the screening[123]. With different FOBT technologies, variable levels of sensitivity are obtained, ranging from 5%-99%, while specificity is somewhat higher 65%-99%[124;125]. Such diverse results indicate that the test is highly dependent on optimal sample preparation and analysis. It also gives both a relatively high degree of false positive and false negative findings. Positive predictive value, or precision rate, is the proportion of patients with who are correctly diagnosed. It is the most important measure of a diagnostic method as it reflects the probability that a positive test reflects the underlying condition being tested for, and a high degree of false positives will make the precision rate decrease. On the contrary, negative predictive value is the proportion of patients with negative test results who are correctly diagnosed and in the same manner, a high rate of false negatives will impair this number. A low precision rate results in unnecessary colonoscopies, hence reducing the cost effectiveness of FOBT screening. Also, FOBT will not detect benign precursor lesions[123]. The requirement for better biomarkers is therefore needed from an early detection screening perspective. A marker with both high sensitivity and specificity will not only prevent extra costs due to unnecessary colonoscopies, but may also increase the time-intervals between repeated tests. Finally, optimal biomarkers should also be present in benign precursor stages and thereby increase the time-window in which removing the polyp is considered curative.

Prognostic and predictive markers

Survival is as mentioned associated with tumor phenotype, as patients with MSI tumors generally have improved prognosis compared to patients with MSS tumors[59]. In addition, 5-flurouracil (5-FU) based chemotherapy does not seem to provide survival benefits among

patients with MSI tumors, neither in stage II nor in stage III colon and rectal cancer [126-128]. There are also survival differences within the MSI-group. In order to further improve prognosis and treatment of this group as a whole it is important to identify those MSI tumors with the worse prognosis which will benefit from non-5-FU based chemotherapy in the future.

In 1997, adjuvant chemotherapy was introduced in Norway for a subset of the CRC patients. Today, with some exceptions, patients with stage I or II receive surgery alone which is considered curative. Patients with metastasis to lymph nodes (stage III) receive chemotherapy in addition to surgery (adjuvant chemotherapy). Different regimes of chemotherapy exist, and the most common is 5-FU/leucovorin in combination with oxaliplatin in patients under 75 years of age[129]. Some stage II patients are also offered chemotherapy when an insufficient number of lymph nodes (< 8) are analyzed for presence of cancer cells. In order to better pinpoint which patients within stage II and III that would benefit from chemotherapy optimally designed studies are needed which can result in improved markers. A flow-chart describing present and possible future elements in CRC diagnostics and choice of treatment is illustrated in Figure 12.

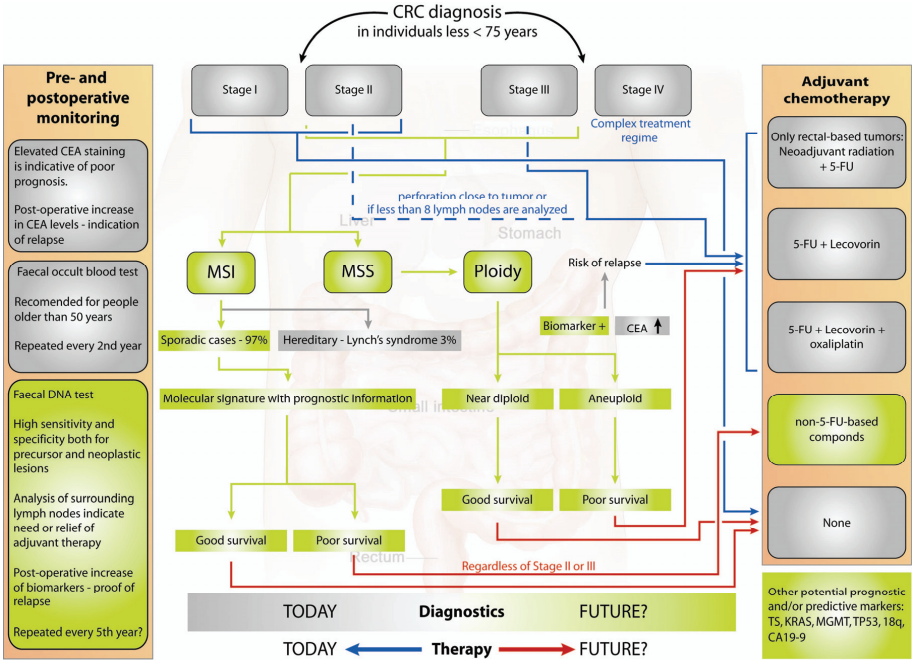


Figure 12. Present and possible future ways of deciding upon CRC treatment. Today stage III receives adjuvant chemotherapy, most usually 5-FU in combination with leucovorin, but other lines of treatment exist. Some rectal tumors receive radiation therapy prior to surgery. Stage I and II are considered to be cured by surgery alone and do not receive additional treatment. Some stage II patients experience recurrence and would probably benefit from adjuvant chemotherapy, whereas a subgroup of stage III individuals will be cured, regardless of chemotherapy. The problem lies in identifying these individuals. The literature contains a large suggestion of predictive and prognostic markers but only CEA (carcinoembryonic antigen) has been implemented in the clinics. However, some of the suggested markers are close to be recommended for clinical practice as they only lack well designed prospective studies verifying their value as markers[130]. Some of these are included in the figure, such as MSI and ploidy. By identifying the stage II patients with a poor prognosis it will be possible to improve prognosis by offering adjuvant chemotherapy.

AIMS

The overall aim of this study is to acquire new biomedical insights in the development of CRC and to use this to pinpoint suitable biomarkers and novel molecular tools for high precision diagnostics and prognosis.

The specific aims were two-fold:

- To explore occurrence of candidate epigenetic biomarkers during colorectal cancer development and their suitability for early tumor detection

- To identify the genetic mutation profiles of selected genes in colorectal carcinomas and their potential use as prognostic markers. These studies included the genes in the MAP kinase pathway and also those affected by mismatch repair deficiency.

RESULTS IN BRIEF

Paper Ia. *"DNA hypermethylation of MAL: a promising diagnostic biomarker for colorectal tumors"*
In a previous study we had identified a short list of candidate epigenetic markers from a genome wide screen[131]. In this commentary to Gastroenterology we reported the initial methylation results for *MAL*, seemingly a highly promising biomarker for early detection of colorectal tumors. *MAL* was frequently methylated both in colorectal carcinomas (83%; 40 of 48 carcinomas) and adenomas (73%; 43 of 59 adenomas) as assessed by methylation-specific polymerase chain reaction (MSP). Hypermethylation of this gene promoter was highly cancer specific as only 11% (2/18) of normal mucosa taken in distance from the primary tumor and 4% (1/23) of normal mucosa from cancer-free individuals were methylated. Methylation frequency in normal mucosa was significantly lower compared to both adenomas and carcinomas ($P < 0.0001$ for both). These findings was in contrast to a recently published article in the same journal, in which *MAL* was reported to be methylated in only 6% of colorectal carcinomas[132]. The authors criticized our results in the commentary and we were not allowed to respond to the response. Thus, the detailed discussion is included in the full length paper (Ib).

Paper Ib. *"Hypermethylated MAL gene – a silent marker of early colon tumorigenesis"*
In the present study, we have compared the promoter methylation status of *MAL* in normal colorectal mucosa samples with benign and malignant colorectal tumors. The article presents the detailed study of *MAL* and includes technical and biological validation experiments. The methylation frequencies were slightly altered from Ia due to addition of more samples. Eighty percent (49/61) of the carcinomas, 71% (45/63) of the adenomas, 10% (2/21) of the normal mucosa from cancer patients, and 4% (1/23) of the normal mucosa from non-cancer individuals harbored hypermethylation of the *MAL* promoter. RNA expression levels of *MAL* were determined for 46 cell lines from various cancer types, and showed that methylation of *MAL* was associated with reduced transcriptional activity ($P = 0.041$). When treating the cell lines with the demethylating agents 5-aza-2'deoxycytidine and trichostatin-A, the cells responded with increased RNA expression, as expected if the transcriptional repression is caused by promoter hypermethylation. The *in situ* protein expression of MAL

was determined by use of a tissue micro array. Among 231 scorable CRC tissue cores, 198 were negative for MAL staining. In conclusion, it seems like *MAL* is inactivated by other mechanisms in addition to hypermethylation as even unmethylated carcinomas showed no protein expression, while positive staining for MAL was seen in all the normal colon mucosa control samples. We speculate that the hypermethylation of *MAL* may act as a "gene expression seal" ensuring that the gene remains in an inactive state.

These findings show that hypermethylated *MAL* is suitable as a diagnostic marker for early colorectal tumorigenesis with a potentially high sensitivity and specificity, the latter calculated between tumors and normal samples. We also show that the low 6% of methylation detected by the Johns Hopkins group was due to suboptimal primer design.

Paper II. *"Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers"*

In this study we analyzed and compared the methylation status of a selected set of markers that potentially could discriminate between non-malignant and malignant tissue from the large bowel. The methylation status of eleven genes (*ADAMTS1*, *CDKN2A*, *CRABP1*, *HOXA9*, *MAL*, *MGMT*, *MLH1*, *NR3C1*, *PTEN*, *RUNX3*, and *SCGB3A1*) was determined by MSP in 154 tissue samples including normal mucosa (20 non-cancerous samples; 18 normal samples from cancer individuals), adenomas (n=63), and carcinomas (n=52) of the colorectum. The reliability of the MSP scorings was tested by quantitative MSP analysis in a blinded manner for one example gene (*MGMT*), and the results were in perfect concordance with the MSP data. Part of the results were previously published[131;133], but was included here in order to analyze co-variance between the different genes. We saw a stepwise, significant increase in methylation frequencies as the mean number of methylated genes per sample was 0.4 in normal colon mucosa from tumor-free individuals, 1.2 in mucosa from cancerous bowels, 2.2 in adenomas, and 3.9 in carcinomas ($P < 0.0001$). This increase in methylation from benign to malignant lesions was also evident at the individual gene level for *ADAMTS1*, *CDKN2A*, *CRABP1*, *MLH1*, *NR3C1*, *RUNX3*, and *SCGB3A1*. We also report that *PTEN* is unmethylated in all carcinomas and is not subject to inactivation by hypermethylation in CRC. Hypermethylation of *CRABP1, MLH1*, *NR3C1*, *RUNX3,* and *SCGB3A1* were seen almost exclusively in proximal carcinomas with microsatellite

instability, and served to classify MSI-tumors. Furthermore, and in agreement with the CIMP concept, the samples with MSI and hypermethylation of several genes also carried *BRAF* mutations. The promoters of *ADAMTS1*, *MAL*, and *MGMT* were frequently methylated both in benign and malignant tumors, independent of microsatellite instability. In addition, *MGMT* was the gene with most frequent promoter hypermethylation among the normal samples taken in distance from primary tumors, which may indicate that it is involved in the creation of a "field effect". From these data we conclude that methylated *ADAMTS1*, *MGMT,* and *MAL* are suitable as markers for early tumor detection.

Paper III. *"Dysregulation of RAS signaling through alterations of RAS, RAF, NF1 and/or RASSF1A in colorectal carcinomas with known microsatellite instability status"*

In this article, four components involved in MAP-kinase signaling were analyzed in a series of colorectal carcinomas. The MAP-kinase pathway is shown to be hyperactive in a large fraction of CRC due to mutations in *KRAS* and *BRAF*, but the possible role of *NF1*, a negative regulator of KRAS signaling, had never been examined in a series of CRC. This is most likely due to the large size of the gene (61 exons) and the fact that no mutation hot-spots or mutation cluster region have been identified. A total of 65 colorectal carcinomas were included in the study, in which mutations in *BRAF* and *KRAS* was assessed by direct sequencing, and promoter hypermethylation of the *RASSF1A* was analyzed by MSP. A representative subset of these tumors (n =24) was included in the *NF1*-analysis, which was performed with denaturing high-performance liquid chromatography (dHPLC), sequencing, multiple ligation-dependent probe amplification (MLPA) and real-time polymerase chain reaction (PCR). Forty percent of the carcinomas (26/65) carried a *KRAS* mutation, and all but two (c.184-189delGAG and c.49insTTG) occurred in codon 12, 13 or 61 which are known to produce a constitutively active protein. *BRAF* was mutated in 22% (14/64) of the carcinomas, and all but three of the mutations were the V600E mutation which also yields a constitutively active protein. Mutations in *KRAS* and *BRAF* were mutually exclusive, and while *BRAF* mutations were strongly associated with MSI (P = 0.006), KRAS mutations were more common among MSS-samples (P = 0.08). We found that 31% (18/59) samples were hypermethylated in the promoter of *RASSF1A*, but there were no covariance between *RASSF1A* methylation and mutation status of either of the analyzed genes. One of the 24 carcinomas analyzed for *NF1* mutations contained two missense mutations and additional

nine tumors displayed intronic mutations in close proximity to the intron–exon boundaries. Using MLPA, we found that another 17% (4/24) samples had a gain of parts or of the whole gene, also confirmed with real-time analysis. Furthermore, 8 of 10 samples with exonic or intronic alterations in *NF1* occurred in MSI-positive tumors (P = 0.047), whereas 3 of 4 duplications occurred in MSS tumors.

In total we found that 74% (48/65) of the tumors most likely had an overactive RAS signaling pathway due to molecular changes of at least one of the four analyzed components. In this study, we found the *NF1* mutation profile to be in contrast both to published germline mutation profile of NF1 patients as well as to the somatic mutation profiles of malignant peripheral nerve sheath tumor taken from patients with and without the NF1 disease[134-136][§§]. One may speculate whether alternative splicing of NF1 is involved in colorectal tumorigenesis.

Paper IV. *"Identification of RCC2 as a prognostic marker among multiple gene mutations in colorectal cancer with defect mismatch repair"*

Forty-one known genes with coding oligonucleotide repeats were analyzed in two series of microsatellite unstable colorectal carcinomas (n = 202) in order to identify frameshift mutations. In a previous literature survey 162 analyzed genes were recognized. Different selection criteria narrowed down the number of genes with a potential impact on tumor development to 41 which were analyzed with fragment analysis. The aim of the study was to subclassify the MSI-tumor into those with good and those with poor prognosis based on mutation profile of one or a combination of the genes.

In total, the two series of MSI-tumors carried a median number of 17 and 19 mutations. A strong association was seen between low mutation frequency and rectal location for individual genes (*ACVR2A*, *ASTE1*, *CASP5*, *MARCKS*, *MBD4*, *MRE11A*, *MSH3*, *TAF1B* and *TFGBR2)* as well as on the total level of mutations (P = 0.008). A big difference in mutation frequency between a small number of MSI-L tumors and the MSI-H tumors was also seen, indicating that a low degree of MSI is insufficient to induce the mutator phenotype in CRC.

[§§] NF1 International Mutation Database (http://www.nfmutation.org)

Univariate survival analysis indicated that several genes could aid in discriminating good and poor prognosis, but only mutations in *RCC2* was associated with a beneficial five-year disease-free survival in both tumor series (P = 0.035 and 0.011, respectively). This finding was confirmed using multivariate analyses even with the inclusion of the strongest known predictor of prognosis to date, tumor stage at diagnosis (P = 0.028 and 0.021, respectively) as mutations in *RCC2* separated patients with a localized disease into those with poor and good survival (P = 0.004)

In conclusion, analysis of an (A)10 repeat in *RCC2* using readily available technology refines prognosis in a group of microsatellite instable tumors.

DISCUSSION

Fresh-frozen versus formalin embedded tissue

In the present study both fresh frozen and formalin-fixed paraffin embedded tissue samples have been used. Formalin fixation is an excellent way of preserving good histological details in tissue sections, and is for this purpose superior to frozen sections. On the other hand, this preservation method is inferior to snap- or fresh frozen tissue for retaining high quality DNA and RNA. One of the main hurdles is the fragmentation of DNA as well as the inhibitory effects on PCR efficiency. This means that one should aspire to design short PCR-products, and carefully optimize the reactions. In the present thesis we only analyzed short PCR products with fragment analysis in the formalin embedded series, and thereby this problem was in large overcome. However, we did experience a small decrease in the success rate between the mutation results obtained from the fresh frozen test series versus those from the formalin-fixed validation series (99.9%; range 99.1-100%, and 92%; range 0-100%, respectively).

Methodological considerations

DNA methylation analyses

Bisulfite treatment

Bisulfite treatment has been around for a while[137], but it was not until the early 1990s that the method was used to map 5-methyl cytosine[138;139]. With this, Frommer and Clarke found a way of converting the non-readable epigenetic information into readable genetic information. Under acidic pH and high bisulfite concentration, bisulfite treatment converts unmethylated, but not methylated, cytosine into uracil in a highly specific manner[140]. The subsequent difference in nucleotide sequence can be exploited to determine methylation status by numerous PCR-based methods. Insufficient bisulfite conversion fails to convert all unmethylated cytosines to thymine, making it difficult to discriminate methylated from unmethylated cytosines in downstream analyses. The conversion rate can be limited by

several factors, and fully denatured DNA, freshly prepared bisulfite solution, low pH for the sulfonation and deamination steps, high pH for the desulfonation reaction, and a free radical scavenger to minimize oxidative damages are essential (described in [140]). These factors are usually well taken care of in the many commercially available kits, including the EpiTect Bisulfite Kit from QIAGEN which was used in the present thesis. This kit results in at least 99% conversion and the repeated denaturation steps during the bisulfite incubation increases the reaction efficiency and significantly reduces the overall time consumption. The protocol is additionally shortened and standardized by clean-up using a QiaCube automated pipetting system.

Methylation specific PCR (MSP)

With MSP, first described in 1996, came the possibility to analyze many genes in larger sample series[141]. The method relies on primer binding specificity to bisulfite converted DNA as one primer pair specifically amplifies methylated DNA and one amplifies unmethylated DNA. Primer design is therefore of crucial importance for a successful and reliable MSP result. Inclusion of multiple CpG sites improves the discrimination of methylated and unmethylated sequences. Non-CpG cytosines should additionally be included in the primers to avoid amplification of any unsuccessfully converted DNA. Without these cytosines, the methylated primers could also amplify unconverted, unmethylated DNA, yielding false positives.

For all MSP primer sets included in papers Ia, Ib and II (expect for *CRABP1*) a minimum of two CpG sites are included in both the sense and antisense primer, ensuring good discrimination between methylated and unmethylated template. For *CRABP1* a single CpG site was included in the sense primer. However, since this was located in the very 3' end of the primer, high specificity was ensured. This was evident from validation analyses where the methylated *CRABP1* primer set was challenged with an unmethylated bisulfite template from normal blood of a healthy person. The reaction produced no product. This was also the case for the rest of the MSP primer pairs used in the present thesis. Additionally, none of the primers amplified unconverted DNA, ensuring that the methylation status from the MSP analyses was not over-estimated.

Since MSP only gives information on the methylation status of the gene in question based on the CpG-sites covered by the primers, it is important that these specific sites are representative for the gene promoter. In cases where you are looking for promoter methylation that affects the gene transcription, CpG sites in the close proximity to the transcription start point will usually be most suitable. Hypermethylation of such promoter regions have in general shown a strong association with loss of or reduced gene expression. In contrast, hypermethylation within the body of the gene may be associated with an active transcriptional state[142]. In the case of *MLH1* only a few bases are correlated with transcriptional repression, while the other CpG sites play a marginal role[143]. Even though a similar result was seen for the *MAL* gene (see bislufite sequencing page 46), both cases represent exceptions to the general link between promoter methylation and transcription. At the same time they underline the importance of good study design and careful technical validation.

Scoring of MSP results

The MSP products were separated with agarose gel electrophoresis, stained with ethidium bromide and visualized with UV. Samples with equal or stronger band intensity than the positive control in the methylation specific reaction were denoted strongly methylated (++), while samples with less intense bands than the positive control were categorized as weakly methylated. Samples with very weak band intensity and those with no visible PCR product in the methylation-specific reaction were regarded as unmethylated. We considered carcinomas with strong band intensities (++) as methylation positive for the gene promoter in question, while the benign lesions and normal mucosa are scored as positive also when weakly methylated (+ and ++). The rationale for this is based on the clonal expansion theory (page 12). In a carcinoma, tumorigenic alterations have already accumulated. Hence, only genes methylated in the majority of tumor cells (++) would be functionally/biologically interesting. In contrast, benign precursors may need additional alterations and/or time in order to develop a malignant potential. Therefore, methylation changes present only in a small fraction of the adenoma cells may provide these cells with a future growth advantage leading to a selection and clonal expansion. By scoring the normal samples and early lesions with a

low threshold, we increase the likelihood of also identifying such changes that are present only in a minor fraction of the sample.

In a diagnostic perspective, all tumor-specific alterations are potentially useful for early detection, regardless of their impact on tumor aggressiveness. In this sense, carcinoma samples could also be scored as positive when weakly methylated. However, the current scoring thresholds provides very conservative results, and ensures that DNA methylation is neither over-estimated in carcinoma samples not under-estimated in normal samples. This also accounts for the resulting sensitivity and specificity measurements.

MSP is a highly sensitive method as 1 methylated allele among 1000 alleles is detectable[141]. Even though the MSP in itself is qualitative and the scoring of the MSP is visual, the results are highly reproducible, which indicates that the band intensities most likely denote the amount of methylated alleles in the lesion. This is verified in paper II, where the *MGMT* gene has been analyzed in a blinded manner with both quantitative and qualitative MSP. Full concordance was achieved, underlining the value of carefully designed assays, well optimized reactions and thorough scoring of the results.

Bisulfite sequencing

Bisulfite sequencing reads out the methylation status of individual CpG sites in a given sequence and is considered to be the most comprehensive method for studying methylation changes. Primers are designed to exclude, or limit, the number of CpG sites in order to amplify both methylated and unmethylated DNA in a non-biased manner. A 5-methyl-cytosine will be read as a cytosine in the final sequence, whereas an unmethylated cytosine will be read as a thymine[139].

Bisulfite sequencing can be performed either directly on the PCR product or by sequencing clones containing the PCR product. The direct sequencing is fast and simple and will provide an average methylation frequency for each of the CpG sites in the entire target sequence. In contrast, the cloning approach is more work demanding, but will provide exact methylation status for each of the CpG sites in the individual clones[144]. The choice of approach

depends on the requirement. Direct sequencing of human tumors can lead to an underestimation of the methylation at CpG sites, since tumor samples often are heterogeneous and may contain normal cells which in most cases are unmethylated for the gene in question.

Bisulfite sequencing is an important validation analysis that will reveal how representative the results from each of the MSP assays are. In Paper Ib, we found hypermethylation of the *MAL* promoter in ~80% of colorectal carcinomas (Paper Ia and Ib) using MSP, while another study (Mori et al) reported only 6% for the same gene[132]. The primers used in our MSP study were located very close to the transcription site while the other study used primers located a couple of hundred base pairs upstream. Bisulfite sequencing clearly showed that methylation is unequally distributed within the *MAL* promoter. A good association between methylation status, as assessed by MSP, and the bisulfite sequences of the overlapping fragment for the MSP product located close to transcription start site (fragment A) was seen. In contrast, the region reported in the Mori study[132] was generally unmethylated in the same cell line panel. Sequence data showed that only a minority of the CpG sites covered by the Mori antisense primer were methylated in the 19 colon cancer cell lines analyzed, despite the fact that these were heavily methylated around the transcription start site. The CpG sites analyzed by the Mori primer set were therefore not representative for the methylation status of the promoter, leading to false negative results and an underestimation of the methylation load of *MAL*. This finding illustrates the importance of combining MSP and bisulfite sequencing when analyzing new genes.

Mutation analyses

Innumerable methods for mutation detection are developed, and the methods can be divided into two crude categories, scanning methods and diagnostic methods. In general, a diagnostic method detects and identifies mutations in one analysis, while a scanning method detects sequence alterations without describing the mutation. Therefore, scanning methods are often combined with a second step that identifies the mutation [18].

It may seem like unnecessary work to perform a scanning method since it must be complemented with a diagnostic method anyway. Why not just go for the diagnostic method and get all results in one analysis? The answer is that different problems require different solutions. If the mutation is already described and we know what we are looking for, the diagnostic approach may be used. The *BRAF*, *KRAS* and *PIK3CA* genes have known mutation hotspots, in which a diagnostic method like direct sequencing is highly suitable. For genes with large coding regions and no known hotspots (exemplified by *NF1*) a scanning method is quite useful. Direct sequencing of the 61 *NF1* exons in 24 carcinomas would require a minimum of ~1500 sequences with a 100% success rate, meaning the final number would be higher. In paper III we pre-screened all 61 exons in 24 samples for *NF1* mutations using denaturing high performance liquid chromatography (DHPLC) and only those with an abnormal elution profile were subjected to direct sequencing. The employment of a scanning approach saved us from performing a large amount of unnecessary sequencing, resulting in a more time and cost effective approach.

Denaturing high performance liquid chromatography

DHPLC was first described in 1995[145] and today several DHPLC-systems are commercially available, such as the MultiMax LH 750 (Rainin Instrument, Woborn, MA, USA) and the WAVE systems (Transgenomic, Crewe, UK). The principle for detection of mutations is temperature-based separation of homoduplex and heteroduplex molecules under partial denaturing conditions[146]. DHPLC detects both single base-pair mismatches and indels of single and multiple bases with success, but are unsuccessful at detecting large genomic rearrangements.

In paper III the WAVE system was used for the DHPLC analysis. Separation of DNA molecules is achieved by means of a mobile phase of hydro-organic eluent containing triethylammonium acetate (TEAA) and acetonitrile, and a solid phase consisting of a column with hydrophobic beads. During the mobile phase the system is first flushed with TEAA, causing the beads to be "coated" with positively charged triethylammonium ion (TEA). TEA is an amphiphilic ion, meaning it has both hydrophobic and hydrophilic ends. This makes the negatively charged DNA molecule able to bind to the positive end of TEA, linking the

DNA fragments to the column. In order to elute the fragments, increasing concentrations of acetonitrile is flushed through the column. As the acetonitrile concentration increases, the bridging capabilities of the TEA ions decrease and the DNA fragments are released. Heteroduplexes, with mismatched base pairs, are eluted first followed by the homoduplexes due to differences in melting temperature. Finally, the fragments pass through an UV detector which detects the absorbance over time at 260nm[146][***].

The sensitivity of DHPLC is by large determined by temperature[147]. Included with the WAVE equipment is the Navigator™ software which based on fragment length and base composition predicts the optimal separation temperature, which is when 75% of the fragment is double stranded. If the temperature is too low, neither the homoduplex nor heteroduplex fragments will be denatured, and will be eluted at the same time. Too high temperature causes all fragments to be fully denatured before acetonitrile is added, again eluting the fragments at the same time. Only at optimal temperatures the heteroduplexes will denature slightly before the homoduplexes, making mutation detection possible. Due to intra-fragment variations in base composition, PCR products longer than 200 bases usually contain two or more melting domains which mean that the product must be analyzed at several temperatures, reducing the throughput of the method. Other factors important for sensitivity are the resolution of the homoduplex and heteroduplex species as well as the purity of the PCR product. With a good study design the sensitivity and specificity of DHPLC is considered to be higher than 96%[146].

In addition to DHPLC, a handful of sensitive mutation scanning methods exist, including SSCP (Single-Strand Conformational Polymorphism)[148], TTGE (temporal temperature gradient gel electrophoresis)[149], CSGE (Conformation Sensitive Gel Electrophoresis)[150] and DGGE (Denaturing Gradient Gel Electrophoresis)[151]. However, most, if not all of these, are clearly unsuitable when analyzing large samples series as these are gel based methods. In this setting a high throughput method would be more suitable, such as conformation-sensitive capillary electrophoresis (CSCE), MALDI-TOF based methods[152], and of course the new generation of ultra high-throughput sequencing systems. Although DHLPC has an inferior throughput compared to these, with an estimated sensitivity higher

[***] Transgenomic web pages – http://www.transgenomic.com

than 96% DHPLC do not stand back in sensitivity for any of these high-throughput methods. In order to be time efficient and not spend time on method optimization, collaboration with a national diagnostic centre for NF1-related diseases in Rome, Italy, was initiated. Here a lab protocol for *NF1* analysis was well established, analyzing both sequence alterations and copy number changes[153;154]. The DHPLC primers were generally positioned approximately 50 to 60 bp away from the intron–exon boundary to allow the detection of splicing defects while minimizing intronic polymorphisms. Melting temperature of all fragments was optimized to yield as high sensitivity as possible. In addition, a large numbers of normal control samples were analyzed and recorded, making it easy to distinguish between mutations and common polymorphisms.

Direct sequencing

Dye terminator sequencing has up until now been considered the golden standard in mutation analysis as it describes any sequence variant with a high accuracy[18]. Still, a limitation with direct sequencing is the detection level. It has been estimated that this method can detect and quantify minor sequence variants mutations present in as little as 10% of a virus population[155], although other studies claim that this number is as high as 30%[18]. The detection level can vary somewhat from sequence to sequence, and although we are able to see base changes at lower cut-offs in a designed sensitivity test using titrations of known mutations, our experience is that the sensitivity level is approximately 15% for unknown mutations. This means that when running several unknown samples, sequence changes will be scored when present at this level or higher. Mutations in genes present in a smaller fraction (than 15%) of the sample (*e.g.* by-stander genes, see page 61-62) will be scored as wild type. In such cases a cloning strategy or another more sensitive method could be used, but one might discuss whether it is interesting, in a biologic perspective, to detect mutations that falls below the detection threshold. Such mutations will per definition only be present in a few of the tumor cells as discussed on pages 61-62, and are not expected to provide the tumor with a selective advantage.

In addition to the inherent nature of random template selection in PCR assays, the base-calling algorithm contributes to this relatively poor sensitivity. Although improvements in

the sequencing technology make it possible to sequence large number of samples, the throughput of direct sequencing is still relatively poor as the data analysis is labor-intensive. Automated sequence scanning softwares are developed, but as with the sensitivity level, the base calling algorithm makes automated mutation calling difficult when the signal to noise ratio is suboptimal. However, with a good signal-to noise ratio, such programs can be very labor-saving when looking at distinct mutation hot-spots such as the V600E in *BRAF*. The program can quickly zoom in on the codon and call possible mutations, and it is easy for the investigator to quickly confirm or refute the results. Hence, genes with well defined mutation regions are more easily adapted to direct sequencing. In the remaining cases manual reading of electropherograms (which requires skills in addition to time) is still preferable.

Here, direct sequencing has been used for analyzing *BRAF* and *KRAS* in a tumor series evaluated to contain an average of 84% tumor cells[156]. Both of these genes contain known mutation hot-spots, and should be easily picked up by an automated analysis. Both genes additionally provide a strong selective advantage for the tumor when mutated and their occurrence is therefore expected to be well above the detection limit of 15%.

Fragment analysis

MSI is also called the mutator phenotype (page 30)[91]. Since the MSI-inflicted indels most often are confined to microsatellite regions, a diagnostic approach is applicable. Direct sequencing could be an option, but due to the previously mentioned specificity issue as well as the workload, other methods such as fragment analysis could be more suitable. This is a sensitive and high throughput method to describe PCR-products when performed with fluorescent primers in a capillary electrophoresis system[157]. Fragment analysis was used when exploring the microsatellite-containing regions of each of 41 genes in paper IV as well as the MSI-analysis. Using this method, the presence and identity of the mutation was easily detected within the same run (Figure 13).

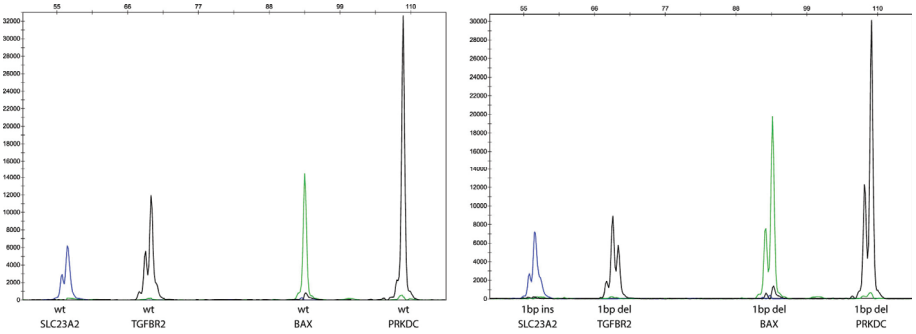


Figure 13. Mutation detection using fragment analysis. The left panel is a sample which is wild-type for all four genes as the electropherograms are identical to the normal control samples. To the right we see a sample with insertion or deletion in all genes. X-axis – fluorescent intensity, Y-axis – fragment size in base pairs.

To date, mutation analyses of close to 200 genes with microsatellites within their coding region have been published in CRC[90] and 41 of the most prominent genes are included in paper IV. With a bioinformatic genome wide approach we have identified that more than 1000 protein coding genes contain 8 or more mononucleotide repeats in their coding sequence[90]. This implies that genes analyzed so far only represent the tip of the iceberg and are not necessarily the ones with most impact on biological or clinical behavior.

Multiplex-dependent probe amplification

As DHPLC fail to detect large genomic rearrangements, another method was included to examine this. MLPA was first described in 2002 as a novel method to detect copy number changes by PCR amplification of several ligated probes[158]. Each MLPA probe consists of two oligonucleotides which can be ligated to each other when hybridized adjacent to each other on a template. All ligated probes have identical 5' and 3' sequences so that they can be amplified with universal PCR primers in a multiplex fashion. Only ligated oligonucleotides can serve as template in a subsequent PCR, eliminating the need for cleaning up unbound probes[159]. Ligated probes are co-amplified and quantified, and a decrease or increase in the amount of the amplified probe indicates loss or gain of the exon, respectively.

Several controls are included in a MLPA assay. Probes located at different genomic locations than the gene of interest are important in order to detect whole gene losses. Control probes scattered around the genome, preferably in regions with known copy number variation can

serve as positive controls. Also, in the case of *NF1*, the exon probes were divided in two separate probe mixes (exon 1, 3, 5 etc. is in one probe mix and exon 2, 4, 6 etc. is in the other). In order to analyze the whole gene, the use of both probe sets will make the quantification of each exon independent from its adjacent exon. In that way there are two experiments that will confirm a whole gene- or a multi exon loss/gain, hence, the two probe sets function as internal controls.

Any factors that may interfere in probe binding will lower MLPA success. Point mutations or indels at, or close to, the ligation site significantly causes reduced probe-sequence affinity. This may cause reduced levels of ligated and amplified product, and the probe to be scored as deleted[159]. In our case this problem was surpassed as we performed mutation analyses in parallel, ensuring that all probes scored as deleted did not contain such aberrations in the proximity to the primers. Striving to have as equal hybridization and amplification efficiency as possible is important as a more effective amplification of one probe may cause it to be mistakenly scored as a gain. Size is also an issue with MLPA as the fragments are separated by this. Theoretically, as smaller fragments are amplified more efficiently than larger, it might seem like a bad choice to use in a multiplex setup. However, if this was a general problem one would assume that the longest probes would be reported as lost in the majority of the cases. In the analysis of *NF1*, we did not see this effect as longer fragments had similar quantities as the smaller ones. As with all quantitative PCR assays, it is important to measure the product while it is in the linear phase, in which the amount of the products are proportional to the copy number of the sequences. Optimal probe design, PCR-setup as well as control regions are all taken care of when using a commercially available kit.

The advantage of MLPA compared to methods such as fluorescent *in situ* hybridization or array-CGH is the fact that it is far less expensive, less labor intensive, and therefore quicker. Overall, MLPA has proved to be a sensitive, time- and cost-efficient method to detect copy number changes for a wide variety of genes[160-162].

Clinical impact of molecular biology

Early tumor markers

DNA promoter hypermethylation has been shown to occur early in the colorectal tumorigenesis[163], making it possible to detect pre-neoplastic lesions. In contrast to gene mutation analyses where multiple tests are required in order to correctly establish the mutation status of the gene in question, promoter hypermethylation requires only one assay per gene, reducing the workload (the principle is illustrated Figure 14). *APC* and *TP53* are examples of genes in which several exons must be analyzed, while *BRAF* and *KRAS* have mutation hot-spots and require only one analysis. These four genes in addition to BAT26, a marker for microsatellite instability, are included in one of the few currently commercially available mutation tests for colorectal cancer. The sensitivity of this and comparable tests have been shown to fall within 52%-91% (reviewed in [164]). When combined with digital melting curves (a method which improves the level of mutation detection to 0.1%, far better than many conventional assays[165]) and/or DNA integrity assay (measures long DNA stretches, more abundant in patients with a tumor[166]), the overall sensitivity has been shown to be around 90% in stool samples from CRC patients with known mutations[165]. As mutations are highly tumor-specific, the specificity of mutation based test are superior to *e.g.* the detection of occult blood in stool samples.

Expression of cancer-specific mRNA transcripts in colonocytes extracted from feces has also a been explored as a potential method for diagnostic tests[167]. When analyzing a panel of four genes (*MMP7*, *MYBL2*, *PTGS2* and *TP53*) in fecal samples, the resulting sensitivity and the specificity were 58% and 88%, respectively. A challenge using this method is that RNA is easily degraded in feces, and only intact colonocytes will provide suitable mRNA for further analyses. In the cited article, a success rate of 75% was obtained[167], a number inferior to *e.g.* analyses of methylated promoters in stool.

So far, only a handful of epigenetic markers with diagnostic potential have been identified[168-170]. Among the most promising ones is *VIM*, which is present in 73% of fecal samples from individuals with CRC. When combined with a DNA integrity assay *VIM* hypermethylation provides a sensitivity and specificity in stool of 88% and 82%,

respectively[169;171]. Several studies have examined the diagnostic potential of *SFRP2* hypermethylation in stool, and report a sensitivity of 77-90% for carcinomas, 46-62% for adenomas and 33-42% for HPs. Specificity is within the range of 77% to 85% [168;172-174].

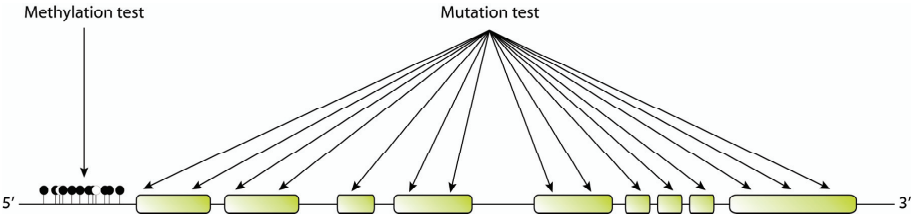


Figure 14. Differences between methylation and mutations tests from a screening perspective. About half the human genes have a CpG-island in its 5' position. Hypermethylation in this region is associated with reduced protein expression, an attribute often exploited by tumor cells. While mutation detection often requires that several exons are examined, and in case of large exons it may require several amplicons, methylation test only need to check the promoter for cancer-specific hypermethylation.

Several DNA methylation markers have also been analyzed in blood samples. This type of starting material is expected to increase patient compliance compared with the alternative non-invasive test, fecal testing. However, with the exception of *SEPT9*, the sensitivity of such markers have generally been low (range 17%-70%)[175-177]. Hypermethylation of *SEPT9* has a sensitivity of 58% and a very high specificity of 90%. By combining the results from two analyses of the same sample; one in normal DNA concentrations and the other in a diluted sample, the sensitivity increased to 72% while the specificity was kept[178]. A drawback with the use of blood plasma as test material is that theoretically, cancer-specific DNA markers are not expected to be shed off the tumor and into the bloodstream in stage I and II patients, who have localized tumors. Still, *SEPT9* methylation has been found also in the bloodstream of patients with large polyps, although with a severely reduced sensitivity (20%). The low success rate in picking up precursor lesions reduces the diagnostic value of analyzing *SEPT9* in blood as people have to develop cancer in order to be test positive.

Another obstacle when analyzing blood samples is the risk of detecting malignancies or diseases in other organs than the large bowel. If the biomarker is not specific enough for colorectal tumors one might end up in a situation in which the patient has a positive finding on a cancer test but no clinically detectable cancer when using colonoscopy. For this purpose

markers with a very high specificity, such as *SEPT9*, are needed. Ideally, tissue specific markers should be included as well. One way of reducing the chance of non-specific positive hits is to use a combination of markers and apply stringent criteria for positive scoring.

In addition to *MAL* (Paper I), five additional markers with similar potential of high sensitivity and specificity are recently identified and validated in colorectal tissue samples in our lab (Lind *et al.*, unpublished). In combination, we may score a tumor as positive if two of the 6 markers are positive. By this we approach 93% of both carcinomas and adenomas are positive and 96% of all normal samples are negative.

Prognostic markers

Tumor features which can predict prognosis are valuable in the clinic. But to date none of the known molecular markers for colorectal cancer are in clinical use, with the exception of monitoring plasmatic CEA levels and genetic testing of known hereditary syndromes[130;179]. One of the major problems with molecular marker studies is that many have limited power, analyzing only small numbers of tumors. Another factor is that the same markers are analyzed with different technologies, which may bias the end results. In spite of this, several markers have shown a prognostic potential, including *APC, TP53* and MSI status. A study has shown that patients presenting tumors with an *APC* mutation before codon 1000 have a shorter cancer-related survival compared to the ones with mutation after this codon[180]. Most *APC*-mutations are truncating, and those that occur before codon 1000 will cause all the ten β-catenin binding sites to be lost. This will most likely result in a stronger deregulation of the WNT-signaling pathway. A similar finding is seen for *TP53* where mutations within the L3 zinc-binding domain are associated with a worse patient prognosis compared to those with mutations outside this domain[181;182]. A large international study has shown that among the twelve possible mutations at codon 12 in *KRAS*, only the G12V mutation is found to have a significant impact on failure-free and overall survival[183]. The molecular marker closest to a prognostic role in the clinic is probably MSI status. A meta-analysis of 32 studies and over 7500 cases confirmed that MSI is significantly associated with good prognosis[59].

Even though the subgroup of MSI-tumors as a whole is associated with a relatively good prognosis, there is still a subgroup within the subgroup with a worse survival. As illustrated in Figure 12, one of our aims is to be able to identify the patients within the MSI-group who are likely to benefit from a more radical form of treatment, possibly by receiving adjuvant chemotherapy. Some studies have analyzed genes prone to indels of the mononucleotide repeats in their coding region in association to patient survival. *ATR* has in a small study been associated with improved survival, although not significant[184]. *TGFBR2* and *BAX* are found to be associated with both poor [185] and good prognosis[186]. The reason for this discrepancy is likely to be caused by small sample sizes (16 and 44 MSI tumors, respectively). Also, the clinical endpoint may affect the results. In the study showing association to poor survival, overall survival was used, while the other study failed to define what kind of survival endpoint they used. Some studies have analyzed protein expression of genes carrying coding mononucleotide repeats and found that low BAX expression is associated with poor survival[187], and that strong staining of RAD50/MRE11/NBS1 was associated with a favorable survival[188]. Both *TGFBR2* and *BAX* have clearly important functions in the cell as they regulate TGF-β signaling (see page 33) and apoptosis, respectively. Even though their involvement in tumorigenesis is obvious, the potential to discriminate those with good and poor prognosis is more uncertain. *TGFBR2* is one of the most frequently mutated genes in MSI-tumors as it is mutated in 96% of right-sided tumors (Paper IV). This will make it impossible to divide a group in two as all tumors have the same alteration. Therefore, the discrimination achieved when comparing survival in those with and without *TGFBR2* mutations are likely to be caused by a difference in tumor location. *BAX* is also more frequently mutated in right-sided tumors, although not significant. Therefore one can not exclude that survival differences associated to *BAX* mutations are due to location rather than mutations as well.

In order to maximize the likelihood of including driver-genes that are likely to have an impact on tumorigenesis and therefore also in discriminating survival, certain selection criteria were employed in Paper IV. Figure 15 shows the size of our study (Paper IV) compared to other studies analyzing coding repetitive units in MSI-tumors. We identified mutations in *RCC2* to be associated with improved prognosis in two independent tumor series of the MSI phenotype. The RCC2 protein is involved in the segregation of

chromosomes in metaphase[189], and loss of expression has been shown to cause G2/M arrest[190]. Indels in the 5'UTR region of this gene, a site of major translational regulation, may affect translational efficiency and stability, leading to cancer cell arrest and reduced proliferation. Knocking out *RCC2* therefore seems like a bad strategy for a cancer cell, and one would expect a strong negative selection pressure. Still, we see a mutation frequency above 50%, indicating that this is not the case.

Samples / Genes	0-5	6-20	21-40	41-70	71-100	101-150	151-250
1-5	8	57	34	9	2	3	2
6-10		6	9	5			
11-15	1		4		1		
16-20		1	1				
21-30			1	1			
31-40							
41-50		1		1			★

Figure 15. Study sizes analyzing coding repetitive units in MSI-tumors. All studies included in the review from 2007[90] are included. Sample size and number of genes analyzed are plotted against each other. Most studies analyze a small number of samples for a restricted number of genes. The star indicates the size of the study presented in paper IV.

Predictive markers

While some biomarkers are well suited for diagnostic purposes, others can be used to determine, or predict, response to a certain therapy. Such markers are called predictive markers. In paper III we showed that mutations in *KRAS* and *BRAF* were common in CRC as they were present in 62% of the tumors. Mutations in these genes have in other studies been shown to predict response to cetuximab, a monoclonal antibody therapy targeting the epidermal growth factor receptor (EGFR). About 30-40% of non-responders to this therapy carry a *KRAS* mutation. Mutations in *BRAF* is significantly associated with lack of treatment response as none of the patients with $BRAF^{mut}$ responded, while none of the responders

were mutated[191]. The reason for this lack of response is that cetuximab targets EGFR, which is located upstream of both KRAS and BRAF. Both these oncogenes are known to become constitutively active when mutated, meaning that they will sustain signaling even in absence of receptor activation[112;192]. In addition to the potential diagnostic use of promoter hypermethylation it may also contribute with predictive information for choice of treatment. This is also the case for MGMT, which is an enzyme involved in direct DNA repair. It works by irreversibly transferring alkyl groups from an O6-guanine to an internal cysteine[193]. Alkylation at the O6-position of guanine is a common point of attack for many carcinogens, and hypermethylated and inactivated *MGMT* makes the cancer cell more susceptible to damage induced by alkylating agents as the damage is left unrepaired. Hypermethylated *MGMT* is associated with prolonged overall and disease-free survival after carmustine treatment compared to the ones without hypermethylation in gliomas[194]. Hence, hypermethylation of *MGMT* serves as a predictive marker for positive response to chemotherapy consisting of alkylating agents. Even though no such agents are used in the standard treatment regime of CRC, a study has shown that hypermethylation of *MGMT* may predict non-recurrence after chemotherapy with 5-FU as these patients have a better outcome[195]. Hypermethylation of *RASSF1A*, a gene commonly hypermethylated in CRC, has been associated with a worse response to cisplatin treatment in both germ cell tumors and hepatoblastomas[196;197]. Cisplatin is a platinum-based cytostatica similar to oxaliplatin, which is used for treating stage III and IV CRC patients. It may be that *RASSF1A* exerts the same negative effect on oxaliplatin treatment in CRC patients, but to my knowledge this has yet to be shown.

Although debated, MSI, caused by hypermethylation of *MLH1* in sporadic CRC, is associated to worse response to 5-FU treatment[59;126-128]. One possible explanation of this cytostatic resistance is that the lack of MMR might allow incorporated 5-FU to cause harmful effects to DNA synthesis and replication, but with no recognition by the dysfunctional MMR system and no inhibition of cell growth. On the other hand, an intact MMR system may trigger a cell death program in MSS colorectal tumors treated with 5-FU, making this agent more effective in this subtype of tumors[128].

These examples underline that we have molecular knowledge that can help to determine a more optimal treatment regime in a more individualized way than what is performed today. The problem is that studies seeking to establish these facts all are lacking, preventing the molecular knowledge from being adapted into clinical use[130;179].

Survival analyses

Survival analyses are often used to measure the effect of a certain drug or a certain marker on clinical outcome. The different endpoints are plentiful and the meaning of each can be confusing as the same clinical endpoint can have different definitions. These differences make it complicated to compare results from different studies and highlights the need for more uniform and well-defined definitions of endpoints[198].

Overall survival is maybe the most unambiguous of the endpoints in survival analyses. It is easy to define, simple to measure and straightforward to interpret[199]. However, the fact that all kinds of mortalities are registered as events may lead to misinterpretation of the results as colorectal patients often are of old age and may die of other reasons than cancer within the follow-up period[200]. Disease-specific survival is an attempt to improve this. A problem that presents itself with this type of analysis is that sudden deaths of cancer patients are often registered as cancer-related deaths while the real cause may be different, leading to an overestimation of cancer-related deaths[200]. This problem is circumvented with overall survival, but again, that approach has its limitations in being over simplistic. Another challenge with survival analyses is the requirement/need for long (and partly costly) follow up period. A large study showed with a correlation of 0.89 that 3-year disease-free survival is a good surrogate endpoint to 5-year overall survival as most of the disease-specific events which occurred within the first three years (such as local recurrences and metastases) led to death within 5 years as most patients with a recurrent disease had died[199]. This finding can have a significant impact on development and clinical testing of novel drugs as the time of clinical trails could be reduced to almost the half of its origin, dramatically reducing the expenses.

In paper IV, our aim was to identify markers in MSI-carcinomas that carry prognostic information. This patient group accounts for about 15% of all CRC patients and are associated with an improved disease outcome compared to that of CRC as a whole. Death from cancer can be a crude measurement as risk of relapse from disease is an important factor when evaluating who would benefit from adjuvant chemotherapy. Hence, disease-free survival was chosen as clinical end-point in order to include both death from cancer and relapse.

Driver and passenger (epi)mutations

Mutations and epigenetic alterations are frequent in human tumors, but most likely not all alterations are important for tumor development. The discrimination between driver and passenger, or bystander, mutations has gained increased interest the last decade. The alterations which give the cell a growth advantage will be selected, denoted driver mutations, while the background noise of mutations that confer no clonal growth advantage are termed passenger mutations[201]. When screening large numbers of genes, and especially with the modern ultra high throughput systems, this discrimination is important.

Several methods are applied in order to separate the passengers from the drivers, most often by use of statistical and probability calculations. Simplified one can say that if a mutation occurs more often than what would be expected by chance, it indicates a positive selection while a more seldom occurrence indicates a negative selection pressure[201-203]. A study by Sjöblom and co-workers included 13023 genes, among which they found over a thousand to be mutated in colorectal and breast cancer. Among those, only 189 were considered to be genes with impact on tumorigenesis, with an average of 9 mutated genes per colorectal carcinoma[202].

The concept of driver and passenger mutations is especially important in tumors with MSI, or the mutator phenotype. The mutability of mononucleotide repetitions within MSI-tumors depends primarily on structural factors. Because of the high background of genetic instability it is challenging to establish which of these alterations plays a role in tumorigenesis. Duval and co-workers has suggested two schemes to divide the genes into driver and by-stander

genes. One is based on mutation frequency and suggests that genes with mutation frequencies below 12% are considered likely to have no functional consequence on tumor development, and therefore entitled by-stander genes[204]. The second scheme is to divide genes into four categories: survivor genes, hibernator genes, cooperator genes and transformator genes[205]. Transformator genes confer a positive selection to the cells, have the highest mutation frequencies, and are most likely to be genes driving tumorigenesis[205]. Prior to Paper IV a literature survey was performed[90] and only genes passing certain criteria, amongst others a mutation frequency exceeding the aforementioned 12% were included. As we aimed to identify prognostic markers we wanted ensure that we did not analyze mere background artifacts.

In the field of epigenetics, genome-wide studies have been performed suggesting that 100-600 genes are subject to promoter hypermethylation in a random tumor[206-208]. Most likely, methylation instability, such as CIMP, causes a large number of genes to be hypermethylated, and similar to the MSI-tumors, only some of the affected genes exert tumorigenic effects.

Whether passenger mutations are just relics of some sort of genome instability and completely irrelevant for tumorigenesis or if they apply some minor effect on tumor growth is still questioned. A recent paper argues that for any trait, there will be only a few genes with large phenotypic effects, but many associated genes having small effects, although their combined effect may be substantial[209]. This may be important to keep in mind when describing a phenotype, but whether or not the passenger mutations have phenotypic effects is irrelevant in the case of identifying diagnostic and maybe also prognostic markers, as illustrated by *VIM* and *MAL*. Indeed, subsequent to our identification of *MAL* as a potential biomarker, it has been shown that hypermethylation of *MAL* is significantly associated with improved disease-free survival in gastric cancer[210]. Hence, *MAL* can be relevant both for diagnosing colorectal cancer and determining prognosis in gastric cancer even though its impact on tumorigenesis is unresolved.

Different origins of colorectal cancer contribute to a non-uniform disease

As mentioned earlier there is increasing evidence that CRC is not a uniform disease, and that there are at least two distinct routes causing it, the traditional adenoma-carcinoma pathway and the newer pathway through HPs and sessile serrated adenomas (page 25-26). There are several indications also in the present studies supporting the theory that MSI tumors originate from sessile serrated adenomas and not traditional adenomas. If adenomas gave rise to both MSS and MSI tumors one would expect close to a similar frequency of MSI-adenomas as MSI-carcinomas (~12-15%). In paper II, 63 adenomas were included in which only 2 were MSI (3%). In addition, the majority of sporadic MSI tumors are caused by *MLH1* hypermethylation. None of the adenomas in our study, not even the two with MSI contained hypermethylation of this gene. Also, the 5 genes almost exclusively methylated in MSI-tumors (*CRABP1*, *MLH1*, *NR3C1*, *RUNX3* and *SCGB3A1*) displayed very low, or absent, methylation frequencies in adenomas, supporting the fact that MSI tumors are unlikely to originate from adenomas. In parallel, we have analyzed the same genes as in paper II for promoter hypermethylation in 12 hyperplastic polyps (HPs) as well as determined mutation status of *BRAF* (data not included). The HPs were significantly more frequently methylated than the adenomas, and when comparing HPs and MSI tumors, similar methylation frequencies were observed across all genes. Also, 50% of the HPs carried a V600E mutation in *BRAF*. Only one of the HPs displayed MSI, consistent with the sessile serrated pathway theory in which mutation in *BRAF* is the initiating event while MSI is a late event[71]. Together these data show that the HPs already carry alterations typical for MSI-tumors, while the adenomas do not, hence supporting the aforementioned theory.

CONCLUSIONS

We have identified novel epigenetic changes in colorectal tumors and some are considered as promising diagnostic biomarkers. Hypermethylation of *ADAMTS1*, *MGMT* and especially *MAL* are frequently seen in precursor lesions as well as carcinomas regardless of MSI-status, and rarely in normal mucosa, making them suitable for early detection. In addition, hypermethylation of *CRABP1*, *MLH1*, *NR3C1*, *RUNX3* and *SCGB3A1* were seen almost exclusively among MSI-carcinomas, making them suitable as identifiers of these tumors. Overall, the augmented methylation frequencies from normal mucosa, via adenomas to carcinomas suggest that epigenetic biomarkers may be used for cancer risk assessment.

For the first time we showed that *NF1*, a negative regulator of KRAS, may be involved in colorectal cancer. Through mutation and copy-number analyses we found that more than 40% of the colorectal carcinomas had mutations in the *NF1* gene, which might cause alternative splicing or wrong protein folding. However, further studies are necessary to verify any potential functional consequence.

Including the *NF1* data we show that alterations in one or more of the *KRAS*, *BRAF*, *NF1* and *RASSF1A* genes are found in 74% of the analyzed carcinomas, underlining the importance of the MAPK pathway in colorectal cancer

The mutation status of one gene, *RCC2*, among 41 genes analyzed was shown to carry prognostic information among patients with MSI tumors. Mutation of *RCC2* was associated with good prognosis, even with the inclusion of tumor stage in multivariate regression analysis. *RCC2* as a prognostic marker was identified in a clinical test series and confirmed in an unselected prospective series.

FUTURE PERSPECTIVES

High throughput technology for detection of genetic and epigenetic cancer markers.
We are planning genome-wide strategies that include the use of a next-generation sequencing system such as SOLiD, 454 or Solexa, opening up innumerable possibilities. This is a very powerful high-throughput technology which allows both genetic and epigenetic genome wide analyses. One possibility is to perform chromatin immunoprecipitation (ChIP) followed by submitting the precipitated DNA fragments to genome wide sequencing analyses (SEQ), a so-called ChIP-SEQ experiment. New targets for epigenetic silencing may be found by sequencing all DNA fragments bound to *e.g.* proteins with a methyl binding domain (MBDs). Sequencing all DNA fragments which bind to a certain transcription factor, can be used to map its downstream effects. The ChIP products may also be combined with microarray studies (ChIP-on-chip) which will enable us to study different kinds of regulation at the transcriptome level.

Non-invasive early diagnostics of colorectal cancer
Even though much effort is put into large-scale studies it is still of uttermost importance to have a good pipeline for detailed methylation analyses of resulting target genes in general and especially genes with a diagnostic potential. In order to achieve this, we are establishing two types of quantitative DNA methylation analyses for all new potential biomarkers, real-time PCR based methods and pyrosequencing. Such technology will be a necessary tool to determine and optimize the detection levels of cancer-specific markers in a non-invasive material such as stool or blood. Potential biomarkers will be analyzed with quantitative studies in series of normal mucosa samples, benign precursor lesions and carcinomas. If the methylation markers have high methylation frequencies in carcinomas and/or precursor lesions and are unmethylated in normal samples, a panel of suitable fecal samples will be analyzed in a blinded manner in order to determine the sensitivity and specificity levels. We are currently working on optimizing the fecal DNA extraction protocol in order to produce maximum yield of high quality, human DNA.

Molecular risk assessment

In addition to our protocol for identification of novel methylation markers, an ongoing project is designed to allow for detection of markers rarely found in polyps but frequent in cancer and as such may be used in risk assessment. A parallel study will analyze some of the newly suggested biomarkers (including *MAL*) in a new clinical series obtained through collaboration with Stavanger University Hospital, including ~200 colorectal tumors and an average of 4 lymph nodes from each patient in order to examine the presence of these biomarkers in the lymph nodes. The aim of the study is to identify a subgroup of stage II patients with molecular evidence of lymph node metastases missed by conventional diagnostics. These will be designated high-risk stage II patients and are likely to benefit from adjuvant chemotherapy.

Quantitative and qualitative transcriptomics.

The tumor transcriptome from patients in two prospective series from different time periods are under analyses using in-lab Affymetrix Exon microarrays. The purpose is to compare the gene signatures of stage II and III colorectal carcinomas from a series collected during 1987-89 (n = 100), a period where no CRC patients received adjuvant chemotherapy, with the validation series from 1997-2003 (n = 100) in which stage III patients and some stage II received adjuvant chemotherapy. The two groups of each series are stratified according to known good or poor disease outcome. With this study design ethical issues regarding the different use of treatment arms are avoided, and potentially we may identify those that will benefit from surgery alone as well as those that will benefit from adjuvant treatment. The exon arrays will also provide data for structural changes within individual genes, which may be due to splice variants or fusion genes. Such changes may be cancer specific and in combination with epigenetic markers these may be used for cancer risk assessment after the initial early tumor detection.

The dynamic genetics and epigenetics of large bowel tumorigenesis.

The challenge in analyses of large datasets obtained through high throughput technologies is even more evident when integrating the various sets. However, such large data sets exist and more will come. The biological value of the pinpointed genes and the transfer to clinical applications will depend on the quality of the study design, including the clinical sampling

and follow up. Combining the molecular results with more detailed biology will improve our understanding of the dynamics in genome stability during the development of a malignancy from a precursor lesion in the large bowel. The necessity of parallel functional studies can be exemplified by the last study of this thesis in which a marker is found associated with good prognosis and currently we do not know why.

REFERENCES

1. Breasted J (1930) **The Edwin Smith Surgical Papyrus.** University of Chicago Press, Chicago,

2. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC (2007) **The diploid genome sequence of an individual human.** PLoS Biol 5(10): e254

3. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM (2008) **The complete genome of an individual by massively parallel DNA sequencing.** Nature 452(7189): 872-876

4. Mendel G (1866) **Versuche über Plflanzen-hybriden. Verhandlungen des naturforschenden Vereines in Brünn.** pp 3-47

5. Hershey AD, Chase M (1952) **Independent functions of viral protein and nucleic acid in growth of bacteriophage.** J Gen Physiol 36(1): 39-56

6. Watson JD, Crick FH (1953) **Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.** Nature 171(4356): 737-738

7. Suter CM, Martin DI, Ward RL (2004) **Germline epimutation of MLH1 in individuals with multiple cancers.** Nat Genet 36(5): 497-501

8. Li E (2002) **Chromatin modification and epigenetic reprogramming in mammalian development.** Nat Rev Genet 3(9): 662-673

9. Funk JO (1999) **Cancer cell cycle control.** Anticancer Res 19(6A): 4772-4780

10. Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C (2003) **The significance of unstable chromosomes in colorectal cancer.** Nat Rev Cancer 3(9): 695-701

11. Nowell PC (1976) **The clonal evolution of tumor cell populations.** Science 194(4260): 23-28

12. Hanahan D, Weinberg RA (2000) **The hallmarks of cancer.** Cell 100(1): 57-70

13. Vogelstein B, Kinzler KW (2004) **Cancer genes and the pathways they control.** Nat Med 10(8): 789-799

14. Grady WM, Carethers JM (2008) **Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis.** Gastroenterology 135(4): 1079-1099

15. Jones PA, Baylin SB (2002) **The fundamental role of epigenetic events in cancer.** Nat Rev Genet 3(6): 415-428

16. Feinberg AP, Vogelstein B (1983) **Hypomethylation of ras oncogenes in primary human cancers.** Biochem Biophys Res Commun 111(1): 47-54

17. Esteller M (2008) **Epigenetics in cancer.** N Engl J Med 358(11): 1148-1159

18. Dahl C, Ralfkiaer U, Guldberg P (2006) **Methods for detection of subtle mutations in cancer genomes.** Crit Rev Oncog 12(1-2): 41-74

19. Hiller M, Platzer M (2008) **Widespread and subtle: alternative splicing at short-distance tandem sites.** Trends Genet 24(5): 246-255

20. Darwin CR (1859) **On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.**

21. International Human Genome Sequencing Consortium (2004) **Finishing the euchromatic sequence of the human genome.** Nature 431(7011): 931-945

22. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR (2004) **A census of human cancer genes.** Nat Rev Cancer 4(3): 177-183

23. Hotchkiss RD (1948) **The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography.** J Biol Chem 175(1): 315-332

24. Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) **Molecular basis of base substitution hotspots in Escherichia coli.** Nature 274(5673): 775-780

25. Shi H, Wang MX, Caldwell CW (2007) **CpG islands: their potential as biomarkers for cancer.** Expert Rev Mol Diagn 7(5): 519-531

26. Goll MG, Bestor TH (2005) **Eukaryotic cytosine methyltransferases.** Annu Rev Biochem 74:481-514. 481-514

27. Chan TL, Yuen ST, Kong CK, Chan YW, Chan AS, Ng WF, Tsui WY, Lo MW, Tam WY, Li VS, Leung SY (2006) **Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer.** Nat Genet 38(10): 1178-1183

28. Hitchins MP, Wong JJ, Suthers G, Suter CM, Martin DI, Hawkins NJ, Ward RL (2007) **Inheritance of a cancer-associated MLH1 germ-line epimutation.** N Engl J Med 356(7): 697-705

29. Deng G, Chen A, Pong E, Kim YS (2001) **Methylation in hMLH1 promoter interferes with its binding to transcription factor CBF and inhibits gene expression.** Oncogene 20(48): 7120-7127

30. Kondo Y, Issa JP (2004) **Epigenetic changes in colorectal cancer.** Cancer Metastasis Rev 23(1-2): 29-39

31. Lopez-Serra L, Esteller M (2008) **Proteins that bind methylated DNA and human cancer: reading the wrong words.** Br J Cancer 98(12): 1881-1885

32. Lopez-Serra L, Ballestar E, Ropero S, Setien F, Billard LM, Fraga MF, Lopez-Nieva P, Alaminos M, Guerrero D, Dante R, Esteller M (2008) **Unmasking of epigenetically silenced candidate tumor suppressor genes by removal of methyl-CpG-binding domain proteins.** Oncogene 27(25): 3556-3566

33. Miremadi A, Oestergaard MZ, Pharoah PD, Caldas C (2007) **Cancer genetics of epigenetic genes.** Hum Mol Genet 16 Spec No 1 R28-R49

34. Ropero S, Fraga MF, Ballestar E, Hamelin R, Yamamoto H, Boix-Chornet M, Caballero R, Alaminos M, Setien F, Paz MF, Herranz M, Palacios J, Arango D, Orntoft TF, Aaltonen LA, Schwartz S Jr, Esteller M (2006) **A truncating mutation of HDAC2 in human cancers confers resistance to histone deacetylase inhibition.** Nat Genet 38(5): 566-569

35. Jenuwein T, Allis CD (2001) **Translating the histone code.** Science 293(5532): 1074-1080

36. Felsenfeld G, Groudine M (2003) **Controlling the double helix.** Nature 421(6921): 448-453

37. Haig D, Westoby M (1989) **Parent-specific gene expression and the triploid endosperm.** Am Nat 134 147-155

38. Burd M (2008) **The Haig-Westoby model revisited.** Am Nat 171(3): 400-404

39. Eden S, Constancia M, Hashimshony T, Dean W, Goldstein B, Johnson AC, Keshet I, Reik W, Cedar H (2001) **An upstream repressor element plays a role in Igf2 imprinting.** EMBO J 20(13): 3518-3525

40. Murrell A, Heeson S, Bowden L, Constancia M, Dean W, Kelsey G, Reik W (2001) **An intragenic methylated region in the imprinted Igf2 gene augments transcription.** EMBO Rep 2(12): 1101-1106

41. Esteller M, Herman JG (2002) **Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours.** J Pathol 196(1): 1-7

42. Yoder JA, Walsh CP, Bestor TH (1997) **Cytosine methylation and the ecology of intragenomic parasites.** Trends Genet 13(8): 335-340

43. Feinberg AP, Vogelstein B (1983) **Hypomethylation distinguishes genes of some human cancers from their normal counterparts.** Nature 301(5895): 89-92

44. Knudson AG, Jr. (1971) **Mutation and cancer: statistical study of retinoblastoma.** Proc Natl Acad Sci U S A 68(4): 820-823

45. Greger V, Passarge E, Hopping W, Messmer E, Horsthemke B (1989) **Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma.** Hum Genet 83(2): 155-158

46. Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, Dryja TP (1991) **Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene.** Am J Hum Genet 48(5): 880-888

47. Deng S, Calin GA, Croce CM, Coukos G, Zhang L (2008) **Mechanisms of microRNA deregulation in human cancer.** Cell Cycle 7(17): 2643-2646

48. Guil S, Esteller M (2008) **DNA methylomes, histone codes and miRNAs: Tying it all together.** Int J Biochem Cell Biol 41(1): 87-95

49. Clark SJ (2007) **Action at a distance: epigenetic silencing of large chromosomal regions in carcinogenesis.** Hum Mol Genet 16 Spec No 1 R88-R95

50. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ (2006) **Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band.** Nat Genet 38(5): 540-549

51. Novak P, Jensen T, Oshiro MM, Wozniak RJ, Nouzova M, Watts GS, Klimecki WT, Kim C, Futscher BW (2006) **Epigenetic inactivation of the HOXA gene cluster in breast cancer.** Cancer Res 66(22): 10664-10670

52. Parkin DM, Bray F, Ferlay J, Pisani P (2005) **Global cancer statistics, 2002.** CA Cancer J Clin 55(2): 74-108

53. Weitz J, Koch M, Debus J, Hohler T, Galle PR, Buchler MW (2005) **Colorectal cancer.** Lancet 365(9454): 153-165

54. Ahmed FE (2004) **Effect of diet, life style, and other environmental/chemopreventive factors on colorectal cancer development, and assessment of the risks.** J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 22(2): 91-147

55. Bodmer WF (2006) **Cancer genetics: colorectal cancer as a model.** J Hum Genet 51(5): 391-396

56. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K (2000) **Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.** N Engl J Med 343(2): 78-85

57. Al-Sukhni W, Aronson M, Gallinger S (2008) **Hereditary colorectal cancer syndromes: familial adenomatous polyposis and lynch syndrome.** Surg Clin North Am 88(4): 819-844

58. Lothe RA, Peltomäki P, Meling GI, Aaltonen LA, Nystrom-Lahti M, Pylkkanen L, Heimdal K, Andersen TI, Moller P, Rognum TO, . (1993) **Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history.** Cancer Res 53(24): 5849-5852

59. Popat S, Hubner R, Houlston RS (2005) **Systematic review of microsatellite instability and colorectal cancer prognosis.** J Clin Oncol 23(3): 609-618

60. Walther A, Houlston R, Tomlinson I (2008) **Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis.** Gut 57(7): 941-950

61. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP (1999) **CpG island methylator phenotype in colorectal cancer.** Proc Natl Acad Sci U S A 96(15): 8681-8686

62. Muto T, Bussey HJ, Morson BC (1975) **The evolution of cancer of the colon and rectum.** Cancer 36(6): 2251-2270

63. Fearon ER, Vogelstein B (1990) **A genetic model for colorectal tumorigenesis.** Cell 61(5): 759-767

64. Kinzler KW, Vogelstein B (1996) **Lessons from hereditary colorectal cancer.** Cell 87(2): 159-170

65. Arends JW (2000) **Molecular interactions in the Vogelstein model of colorectal carcinoma.** J Pathol 190(4): 412-416

66. Lane DP (1992) **Cancer. p53, guardian of the genome.** Nature 358(6381): 15-16

67. Jass JR (2006) **Colorectal cancer: a multipathway disease.** Crit Rev Oncog 12(3-4): 273-287

68. Snover DC, Jass JR, Fenoglio-Preiser C, Batts KP (2005) **Serrated polyps of the large intestine: a morphologic and molecular review of an evolving concept.** Am J Clin Pathol 124(3): 380-391

69. Torlakovic E, Snover DC (1996) **Serrated adenomatous polyposis in humans.** Gastroenterology 110(3): 748-755

70. Longacre TA, Fenoglio-Preiser CM (1990) **Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia.** Am J Surg Pathol 14(6): 524-537

71. O'Brien MJ, Yang S, Mack C, Xu H, Huang CS, Mulcahy E, Amorosino M, Farraye FA (2006) **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 30(12): 1491-1501

72. Belshaw NJ, Elliott GO, Foxall RJ, Dainty JR, Pal N, Coupe A, Garg D, Bradburn DM, Mathers JC, Johnson IT (2008) **Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa.** Br J Cancer 99(1): 136-142

73. Ramirez N, Bandres E, Navarro A, Pons A, Jansa S, Moreno I, Martinez-Rodenas F, Zarate R, Bitarte N, Monzo M, Garcia-Foncillas J (2008) **Epigenetic events in normal colonic mucosa surrounding colorectal cancer lesions.** Eur J Cancer 44(17): 2689-2695

74. 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 (2005) **MGMT promoter methylation and field defect in sporadic colorectal cancer.** J Natl Cancer Inst 97(18): 1330-1338

75. Giovannucci E, Ogino S (2005) **DNA methylation, field effects, and colorectal cancer.** J Natl Cancer Inst 97(18): 1317-1319

76. Boveri T (1914) **Zur Frage der Entstehung maligner Tumoren.**

77. Peltomaki P, Aaltonen LA, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Green JS, Jass JR, Weber JL, Leach FS, . (1993) **Genetic mapping of a locus predisposing to human colorectal cancer.** Science 260(5109): 810-812

78. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Powell SM, Jen J, Hamilton SR, . (1993) **Clues to the pathogenesis of familial colorectal cancer.** Science 260(5109): 812-816

79. Thibodeau SN, Bren G, Schaid D (1993) **Microsatellite instability in cancer of the proximal colon.** Science 260(5109): 816-819

80. Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) **Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review.** Mol Ecol 11(12): 2453-2465

81. de la Chapelle A (2003) **Microsatellite instability.** N Engl J Med 349(3): 209-210

82. Kunkel TA (1990) **Misalignment-mediated DNA synthesis errors.** Biochemistry 29(35): 8003-8011

83. Lovett ST (2004) **Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences.** Mol Microbiol 52(5): 1243-1253

84. Kunkel TA (1993) **Nucleotide repeats. Slippery DNA and diseases.** Nature 365(6443): 207-208

85. Loeb LA, Monnat RJ, Jr. (2008) **DNA polymerases and human disease.** Nat Rev Genet 9(8): 594-604

86. Kunkel TA, Bebenek K (2000) **DNA replication fidelity.** Annu Rev Biochem 69:497-529. 497-529

87. Kuismanen SA, Holmberg MT, Salovaara R, de la CA, Peltomäki P (2000) **Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers.** Am J Pathol 156(5): 1773-1779

88. Duval A, Iacopetta B, Thorstensen L, Meling GI, Lothe RA, Thuille B, Suraweera N, Thomas G, Hamelin R (2001) **Gender difference for mismatch repair deficiency in human colorectal cancer.** Gastroenterology 121(4): 1026-1027

89. Kroutil LC, Register K, Bebenek K, Kunkel TA (1996) **Exonucleolytic proofreading during replication of repetitive DNA.** Biochemistry 35(3): 1046-1053

90. Røyrvik EC, Ahlquist T, Rognes T, Lothe RA (2007) **Slip slidin' away: a duodecennial review of targeted genes in mismatch repair deficient colorectal cancer.** Crit Rev Oncog 13(3): 229-257

91. Shibata D, Peinado MA, Ionov Y, Malkhosyan S, Perucho M (1994) **Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation.** Nat Genet 6(3): 273-281

92. Issa JP (2004) **CpG island methylator phenotype in cancer.** Nat Rev Cancer 4(12): 988-993

93. Goel A, Arnold CN, Niedzwiecki D, Chang DK, Ricciardiello L, Carethers JM, Dowell JM, Wasserman L, Compton C, Mayer RJ, Bertagnolli MM, Boland CR (2003) **Characterization of sporadic colon cancer by patterns of genomic instability.** Cancer Res 63(7): 1608-1614

94. Shen L, Catalano PJ, Benson AB, III, O'Dwyer P, Hamilton SR, Issa JP (2007) **Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy.** Clin Cancer Res 13(20): 6093-6098

95. Teodoridis JM, Hardie C, Brown R (2008) **CpG island methylator phenotype (CIMP) in cancer: causes and implications.** Cancer Lett 268(2): 177-186

96. van RM, Elsaleh H, Joseph D, McCaul K, Iacopetta B (2003) **CpG island methylator phenotype is an independent predictor of survival benefit from 5-fluorouracil in stage III colorectal cancer.** Clin Cancer Res 9(8): 2898-2903

97. Samowitz WS, Albertsen H, Herrick J, Levin TR, Sweeney C, Murtaugh MA, Wolff RK, Slattery ML (2005) **Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer.** Gastroenterology 129(3): 837-845

98. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP (2007) **Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer.** Proc Natl Acad Sci U S A 104(47): 18654-18659

99. Samowitz WS, Slattery ML, Sweeney C, Herrick J, Wolff RK, Albertsen H (2007) **APC mutations and other genetic and epigenetic changes in colon cancer.** Mol Cancer Res 5(2): 165-170

100. Yang S, Farraye FA, Mack C, Posnik O, O'Brien MJ (2004) **BRAF and KRAS Mutations in hyperplastic polyps and serrated adenomas of the colorectum: relationship to histology and CpG island methylation status.** Am J Surg Pathol 28(11): 1452-1459

101. van RM, Grieu F, Elsaleh H, Joseph D, Iacopetta B (2002) **Characterisation of colorectal cancers showing hypermethylation at multiple CpG islands.** Gut 51(6): 797-802

102. Whitehall VL, Wynter CV, Walsh MD, Simms LA, Purdie D, Pandeya N, Young J, Meltzer SJ, Leggett BA, Jass JR (2002) **Morphological and molecular heterogeneity within nonmicrosatellite instability-high colorectal cancer.** Cancer Res 62(21): 6011-6014

103. Anacleto C, Leopoldino AM, Rossi B, Soares FA, Lopes A, Rocha JC, Caballero O, Camargo AA, Simpson AJ, Pena SD (2005) **Colorectal cancer "methylator phenotype": fact or artifact?** Neoplasia 7(4): 331-335

104. Lee S, Hwang KS, Lee HJ, Kim JS, Kang GH (2004) **Aberrant CpG island hypermethylation of multiple genes in colorectal neoplasia.** Lab Invest 84(7): 884-893

105. Yamashita K, Dai T, Dai Y, Yamamoto F, Perucho M (2003) **Genetics supersedes epigenetics in colon cancer phenotype.** Cancer Cell 4(2): 121-131

106. Issa JP, Shen L, Toyota M (2005) **CIMP, at last.** Gastroenterology 129(3): 1121-1124

107. Goel A, Nagasaka T, Arnold CN, Inoue T, Hamilton C, Niedzwiecki D, Compton C, Mayer RJ, Goldberg R, Bertagnolli MM, Boland CR (2007) **The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer.** Gastroenterology 132(1): 127-138

108. 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 (2006) **CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer.** Nat Genet 38(7): 787-793

109. Lee S, Cho NY, Yoo EJ, Kim JH, Kang GH (2008) **CpG island methylator phenotype in colorectal cancers: comparison of the new and classic CpG island methylator phenotype marker panels.** Arch Pathol Lab Med 132(10): 1657-1665

110. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S (1998) **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 58(22): 5248-5257

111. Ehrlich M (2003) **The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease.** Clin Immunol 109(1): 17-28

112. Dhillon AS, Hagan S, Rath O, Kolch W (2007) **MAP kinase signalling pathways in cancer.** Oncogene 26(22): 3279-3290

113. Oving IM, Clevers HC (2002) **Molecular causes of colon cancer.** Eur J Clin Invest 32(6): 448-457

114. Thorstensen L, Lothe RA (2003) **The WNT Signaling Pathway and Its Role in Human Solid Tumors.** Atlas Genet Cytogenet Oncol Haematol. http://atlasgeneticsoncology.org/deep/WNTSignPathID20042.html

115. Logan CY, Nusse R (2004) **The Wnt signaling pathway in development and disease.** Annu Rev Cell Dev Biol 20:781-810. 781-810

116. Fang JY, Richardson BC (2005) **The MAPK signalling pathways and colorectal cancer.** Lancet Oncol 6(5): 322-327

117. Engelman JA, Luo J, Cantley LC (2006) **The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism.** Nat Rev Genet 7(8): 606-619

118. Shaw RJ, Cantley LC (2006) **Ras, PI(3)K and mTOR signalling controls tumour cell growth.** Nature 441(7092): 424-430

119. Massague J (2008) **TGFbeta in Cancer.** Cell 134(2): 215-230

120. Xu Y, Pasche B (2007) **TGF-beta signaling alterations and susceptibility to colorectal cancer.** Hum Mol Genet 16 Spec No 1 R14-R20

121. Derynck R, Akhurst RJ (2007) **Differentiation plasticity regulated by TGF-beta family proteins in development and disease.** Nat Cell Biol 9(9): 1000-1004

122. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH (1999) **Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I.** Scand J Gastroenterol 34(4): 414-420

123. Bretthauer M, Hoff G (2007) **[Prevention and early diagnosis of colorectal cancer].** Tidsskr Nor Laegeforen 127(20): 2688-2691

124. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E (2007) **Screening for colorectal cancer using the faecal occult blood test, Hemoccult.** Cochrane Database Syst Rev(1): CD001216

125. Burch JA, Soares-Weiser K, St John DJ, Duffy S, Smith S, Kleijnen J, Westwood M (2007) **Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: a systematic review.** J Med Screen 14(3): 132-137

126. Carethers JM, Smith EJ, Behling CA, Nguyen L, Tajima A, Doctolero RT, Cabrera BL, Goel A, Arnold CA, Miyai K, Boland CR (2004) **Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer.** Gastroenterology 126(2): 394-401

127. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, Hamilton SR, Laurent-Puig P, Gryfe R, Shepherd LE, Tu D, Redston M, Gallinger S (2003) **Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer.** N Engl J Med 349(3): 247-257

128. Benatti P, Gafa R, Barana D, Marino M, Scarselli A, Pedroni M, Maestri I, Guerzoni L, Roncucci L, Menigatti M, Roncari B, Maffei S, Rossi G, Ponti G, Santini A, Losi L, Di GC, Oliani C, Ponz de LM, Lanza G (2005) **Microsatellite instability and colorectal cancer prognosis.** Clin Cancer Res 11(23): 8332-8340

129. Dahl O (2007) **[Adjuvant chemotherapy for colon cancer].** Tidsskr Nor Laegeforen 127(23): 3094-3096

130. Duffy MJ, van DA, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, Lamerz R, Peltomaki P, Sturgeon C, Topolcan O (2007) **Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use.** Eur J Cancer 43(9): 1348-1360

131. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, Lothe RA (2006) **ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis.** Cell Oncol 28(5-6): 259-272

132. Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM, Meltzer SJ (2006) **A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer.** Gastroenterology 131(3): 797-808

133. Lind GE, Thorstensen L, Løvig T, Meling GI, Hamelin R, Rognum TO, Esteller M, Lothe RA (2004) **A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines.** Mol Cancer 3:28

134. De Luca A, Buccino A, Gianni D, Mangino M, Giustini S, Richetta A, Divona L, Calvieri S, Mingarelli R, Dallapiccola B (2003) **NF1 gene analysis based on DHPLC.** Hum Mutat 21(2): 171-172

135. Upadhyaya M, Kluwe L, Spurlock G, Monem B, Majounie E, Mantripragada K, Ruggieri M, Chuzhanova N, Evans DG, Ferner R, Thomas N, Guha A, Mautner V (2007) **Germline and somatic NF1 gene mutation spectrum in NF1-associated malignant peripheral nerve sheath tumors (MPNSTs).** Hum Mutat 29(1): 74-82

136. Bottillo I, Ahlquist T, Brekke HR, Danielsen SA, van den Berg E, Mertens F, Lothe RA, Dallapiccola B. **Germline and somatic NF1 mutations in sporadic and NF1-associated malignant peripheral nerve sheath tumors.** The Journal of Pathology. *In Press*

137. Shapiro R, Weisgras JM (1970) **Bisulfite-catalyzed transamination of cytosine and cytidine.** Biochem Biophys Res Commun 40(4): 839-843

138. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL (1992) **A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.** Proc Natl Acad Sci U S A 89(5): 1827-1831

139. Clark SJ, Harrison J, Paul CL, Frommer M (1994) **High sensitivity mapping of methylated cytosines.** Nucleic Acids Res 22(15): 2990-2997

140. Clark SJ (2005) Studying Mammalian DNA Methylation: Bisulfite Modification. In: Esteller M (ed) **DNA Methylation - Approaches, Methods, and Applications**. CRC press, pp 53-63

141. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB (1996) **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.** Proc Natl Acad Sci U S A 93(18): 9821-9826

142. Jones PA (1999) **The DNA methylation paradox.** Trends Genet 15(1): 34-37

143. Deng G, Chen A, Hong J, Chae HS, Kim YS (1999) **Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression.** Cancer Res 59(9): 2029-2033

144. Herman JG (2005) Bisulfite PCR-Based Techniques for the Study of DNA Methylation. In: Esteller M (ed) **DNA Methylation - Approaches, Methods, and Applications**. CRC press, pp 65-72

145. Oefner PJ, Underhill P (1995) **Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC).** Am J Hum Genet 57(A266):

146. Xiao W, Oefner PJ (2001) **Denaturing high-performance liquid chromatography: A review.** Hum Mutat 17(6): 439-474

147. Jones AC, Austin J, Hansen N, Hoogendoorn B, Oefner PJ, Cheadle JP, O'Donovan MC (1999) **Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis.** Clin Chem 45(8 Pt 1): 1133-1140

148. Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) **Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction.** Genomics 5(4): 874-879

149. Sørlie T, Johnsen H, Vu P, Lind GE, Lothe RA, Børresen-Dale AL (2005) **Mutation screening of the TP53 gene by temporal temperature gradient gel electrophoresis.** Methods Mol Biol 291 207-216

150. Ganguly A, Rock MJ, Prockop DJ (1993) **Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes.** Proc Natl Acad Sci U S A 90(21): 10325-10329

151. Fischer SG, Lerman LS (1983) **DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory.** Proc Natl Acad Sci U S A 80(6): 1579-1583

152. Stanssens P, Zabeau M, Meersseman G, Remes G, Gansemans Y, Storm N, Hartmer R, Honisch C, Rodi CP, Bocker S, van den BD (2004) **High-throughput MALDI-TOF discovery of genomic sequence polymorphisms.** Genome Res 14(1): 126-133

153. De Luca A, Schirinzi A, Buccino A, Bottillo I, Sinibaldi L, Torrente I, Ciavarella A, Dottorini T, Porciello R, Giustini S, Calvieri S, Dallapiccola B (2004) **Novel and recurrent mutations in the NF1 gene in Italian patients with neurofibromatosis type 1.** Hum Mutat 23(6): 629

154. De Luca A, Bottillo I, Dasdia MC, Morella A, Lanari V, Bernardini L, Divona L, Giustini S, Sinibaldi L, Novelli A, Torrente I, Schirinzi A, Dallapiccola B (2007) **Deletions of NF1 gene and exons detected by multiplex ligation-dependent probe amplification.** J Med Genet 44(12): 800-808

155. Leitner T, Halapi E, Scarlatti G, Rossi P, Albert J, Fenyo EM, Uhlen M (1993) **Analysis of heterogeneous viral populations by direct DNA sequencing.** Biotechniques 15(1): 120-127

156. Meling GI, Lothe RA, Børresen AL, Hauge S, Graue C, Clausen OP, Rognum TO (1991) **Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis.** Br J Cancer 64(3): 475-480

157. Berg KD, Glaser CL, Thompson RE, Hamilton SR, Griffin CA, Eshleman JR (2000) **Detection of microsatellite instability by fluorescence multiplex polymerase chain reaction.** J Mol Diagn 2(1): 20-28

158. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G (2002) **Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification.** Nucleic Acids Res 30(12): e57

159. Den Dunnen JT, White SJ (2006) **MLPA and MAPH: sensitive detection of deletions and duplications.** Curr Protoc Hum Genet Chapter 7:Unit 7.14.

160. Vorstman JA, Jalali GR, Rappaport EF, Hacker AM, Scott C, Emanuel BS (2006) **MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q.** Hum Mutat 27(8): 814-821

161. Hogervorst FB, Nederlof PM, Gille JJ, McElgunn CJ, Grippeling M, Pruntel R, Regnerus R, van WT, van SR, Menko FH, Kluijt I, Dommering C, Verhoef S, Schouten JP, van't Veer LJ, Pals G (2003) **Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method.** Cancer Res 63(7): 1449-1453

162. Schwartz M, Duno M (2005) **Multiplex ligation-dependent probe amplification is superior for detecting deletions/duplications in Duchenne muscular dystrophy.** Clin Genet 67(2): 189-191

163. Chan AO, Broaddus RR, Houlihan PS, Issa JP, Hamilton SR, Rashid A (2002) **CpG island methylation in aberrant crypt foci of the colorectum.** Am J Pathol 160(5): 1823-1830

164. Loganayagam A (2008) **Faecal screening of colorectal cancer.** Int J Clin Pract 62(3): 454-459

165. Zou H, Taylor WR, Harrington JJ, Hussain FT, Cao X, Loprinzi CL, Levine TR, Rex DK, Ahnen D, Knigge KL, Lance P, Jiang X, Smith DI, Ahlquist DA (2008) **High Detection Rates of Colorectal Neoplasia by Stool DNA Testing With a Novel Digital Melt Curve Assay.** Gastroenterology

166. Syngal S, Stoffel E, Chung D, Willett C, Schoetz D, Schroy P, Jagadeesh D, Morel K, Ross M (2006) **Detection of stool DNA mutations before and after treatment of colorectal neoplasia.** Cancer 106(2): 277-283

167. Koga Y, Yasunaga M, Moriya Y, Akasu T, Fujita S, Yamamoto S, Kozu T, Baba H, Matsumura Y (2008) **Detection of colorectal cancer cells from feces using**

quantitative real-time RT-PCR for colorectal cancer diagnosis. Cancer Sci 99(10): 1977-1983

168. Muller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, Muhlthaler M, Ofner D, Margreiter R, Widschwendter M (2004) **Methylation changes in faecal DNA: a marker for colorectal cancer screening?** Lancet 363(9417): 1283-1285

169. 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 (2005) **Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene.** J Natl Cancer Inst 97(15): 1124-1132

170. 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 (2008) **DNA methylation biomarkers for blood-based colorectal cancer screening.** Clin Chem 54(2): 414-423

171. Itzkowitz SH, Jandorf L, Brand R, Rabeneck L, Schroy PC, III, Sontag S, Johnson D, Skoletsky J, Durkee K, Markowitz S, Shuber A (2007) **Improved fecal DNA test for colorectal cancer screening.** Clin Gastroenterol Hepatol 5(1): 111-117

172. Oberwalder M, Zitt M, Wontner C, Fiegl H, Goebel G, Zitt M, Kohle O, Muhlmann G, Ofner D, Margreiter R, Muller HM (2008) **SFRP2 methylation in fecal DNA--a marker for colorectal polyps.** Int J Colorectal Dis 23(1): 15-19

173. Wang DR, Tang D (2008) **Hypermethylated SFRP2 gene in fecal DNA is a high potential biomarker for colorectal cancer noninvasive screening.** World J Gastroenterol 14(4): 524-531

174. Huang Z, Li L, Wang J (2007) **Hypermethylation of SFRP2 as a potential marker for stool-based detection of colorectal cancer and precancerous lesions.** Dig Dis Sci 52(9): 2287-2291

175. Tan SH, Ida H, Lau QC, Goh BC, Chieng WS, Loh M, Ito Y (2007) **Detection of promoter hypermethylation in serum samples of cancer patients by methylation-specific polymerase chain reaction for tumour suppressor genes including RUNX3.** Oncol Rep 18(5): 1225-1230

176. Nakayama G, Hibi K, Nakayama H, Kodera Y, Ito K, Akiyama S, Nakao A (2007) **A highly sensitive method for the detection of p16 methylation in the serum of colorectal cancer patients.** Anticancer Res 27(3B): 1459-1463

177. Leung WK, To KF, Man EP, Chan MW, Bai AH, Hui AJ, Chan FK, Sung JJ (2005) **Quantitative detection of promoter hypermethylation in multiple genes in the serum of patients with colorectal cancer.** Am J Gastroenterol 100(10): 2274-2279

178. Grutzmann R, Molnar B, Pilarsky C, Habermann JK, Schlag PM, Saeger HD, Miehlke S, Stolz T, Model F, Roblick UJ, Bruch HP, Koch R, Liebenberg V, Devos

T, Song X, Day RH, Sledziewski AZ, Lofton-Day C (2008) **Sensitive Detection of Colorectal Cancer in Peripheral Blood by Septin 9 DNA Methylation Assay.** PLoS ONE 3(11): e3759

179. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC, Jr. (2006) **ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer.** J Clin Oncol 24(33): 5313-5327

180. Løvig T, Meling GI, Diep CB, Thorstensen L, Norheim AS, Lothe RA, Rognum TO (2002) **APC and CTNNB1 mutations in a large series of sporadic colorectal carcinomas stratified by the microsatellite instability status.** Scand J Gastroenterol 37(10): 1184-1193

181. Diep CB, Thorstensen L, Meling GI, Skovlund E, Rognum TO, Lothe RA (2003) **Genetic tumor markers with prognostic impact in Dukes' stages B and C colorectal cancer patients.** J Clin Oncol 21(5): 820-829

182. Børresen-Dale AL, Lothe RA, Meling GI, Hainaut P, Rognum TO, Skovlund E (1998) **TP53 and long-term prognosis in colorectal cancer: mutations in the L3 zinc-binding domain predict poor survival.** Clin Cancer Res 4(1): 203-210

183. Andreyev HJ, Norman AR, Cunningham D, Oates J, Dix BR, Iacopetta BJ, Young J, Walsh T, Ward R, Hawkins N, Beranek M, Jandik P, Benamouzig R, Jullian E, Laurent-Puig P, Olschwang S, Muller O, Hoffmann I, Rabes HM, Zietz C, Troungos C, Valavanis C, Yuen ST, Ho JW, Croke CT, O'Donoghue DP, Giaretti W, Rapallo A, Russo A, Bazan V, Tanaka M, Omura K, Azuma T, Ohkusa T, Fujimori T, Ono Y, Pauly M, Faber C, Glaesener R, de Goeij AF, Arends JW, Andersen SN, Løvig T, Breivik J, Gaudernack G, Clausen OP, De Angelis PD, Meling GI, Rognum TO, Smith R, Goh HS, Font A, Rosell R, Sun XF, Zhang H, Benhattar J, Losi L, Lee JQ, Wang ST, Clarke PA, Bell S, Quirke P, Bubb VJ, Piris J, Cruickshank NR, Morton D, Fox JC, Al-Mulla F, Lees N, Hall CN, Snary D, Wilkinson K, Dillon D, Costa J, Pricolo VE, Finkelstein SD, Thebo JS, Senagore AJ, Halter SA, Wadler S, Malik S, Krtolica K, Urosevic N (2001) **Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study.** Br J Cancer 85(5): 692-696

184. Lewis KA, Bakkum-Gamez J, Loewen R, French AJ, Thibodeau SN, Cliby WA (2007) **Mutations in the ataxia telangiectasia and rad3-related-checkpoint kinase 1 DNA damage response axis in colon cancers.** Genes Chromosomes Cancer 46(12): 1061-1068

185. Fernandez-Peralta AM, Nejda N, Oliart S, Medina V, Azcoita MM, Gonzalez-Aguilera JJ (2005) **Significance of mutations in TGFBR2 and BAX in neoplastic progression and patient outcome in sporadic colorectal tumors with high-frequency microsatellite instability.** Cancer Genet Cytogenet 157(1): 18-24

186. Jung B, Smith EJ, Doctolero RT, Gervaz P, Alonso JC, Miyai K, Keku T, Sandler RS, Carethers JM (2006) **Influence of target gene mutations on survival, stage**

and histology in sporadic microsatellite unstable colon cancers. Int J Cancer 118(10): 2509-2513

187. Nehls O, Okech T, Hsieh CJ, Enzinger T, Sarbia M, Borchard F, Gruenagel HH, Gaco V, Hass HG, Arkenau HT, Hartmann JT, Porschen R, Gregor M, Klump B (2007) **Studies on p53, BAX and Bcl-2 protein expression and microsatellite instability in stage III (UICC) colon cancer treated by adjuvant chemotherapy: major prognostic impact of proapoptotic BAX.** Br J Cancer 96(9): 1409-1418

188. Gao J, Zhang H, Arbman G, Sun XF (2008) **RAD50/MRE11/NBS1 proteins in relation to tumour development and prognosis in patients with microsatellite stable colorectal cancer.** Histol Histopathol 23(12): 1495-1502

189. Martineau-Thuillier S, Andreassen PR, Margolis RL (1998) **Colocalization of TD-60 and INCENP throughout G2 and mitosis: evidence for their possible interaction in signalling cytokinesis.** Chromosoma 107(6-7): 461-470

190. Mollinari C, Reynaud C, Martineau-Thuillier S, Monier S, Kieffer S, Garin J, Andreassen PR, Boulet A, Goud B, Kleman JP, Margolis RL (2003) **The mammalian passenger protein TD-60 is an RCC1 family member with an essential role in prometaphase to metaphase progression.** Dev Cell 5(2): 295-307

191. Di NF, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P, De DS, Mazzucchelli L, Frattini M, Siena S, Bardelli A (2008) **Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer.** J Clin Oncol 26(35): 5705-5712

192. Lievre A, Bachet JB, Le CD, Boige V, Landi B, Emile JF, Cote JF, Tomasic G, Penna C, Ducreux M, Rougier P, Penault-Llorca F, Laurent-Puig P (2006) **KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer.** Cancer Res 66(8): 3992-3995

193. Pegg AE (1990) **Mammalian O6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents.** Cancer Res 50(19): 6119-6129

194. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG (2000) **Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents.** N Engl J Med 343(19): 1350-1354

195. Nagasaka T, Sharp GB, Notohara K, Kambara T, Sasamoto H, Isozaki H, MacPhee DG, Jass JR, Tanaka N, Matsubara N (2003) **Hypermethylation of O6-methylguanine-DNA methyltransferase promoter may predict nonrecurrence after chemotherapy in colorectal cancer cases.** Clin Cancer Res 9(14): 5306-5312

196. Honda S, Haruta M, Sugawara W, Sasaki F, Ohira M, Matsunaga T, Yamaoka H, Horie H, Ohnuma N, Nakagawara A, Hiyama E, Todo S, Kaneko Y (2008) **The methylation status of RASSF1A promoter predicts responsiveness to chemotherapy and eventual cure in hepatoblastoma patients.** Int J Cancer 123(5): 1117-1125

197. Koul S, McKiernan JM, Narayan G, Houldsworth J, Bacik J, Dobrzynski DL, Assaad AM, Mansukhani M, Reuter VE, Bosl GJ, Chaganti RS, Murty VV (2004) **Role of promoter hypermethylation in Cisplatin treatment response of male germ cell tumors.** Mol Cancer 3 16

198. Punt CJ, Buyse M, Kohne CH, Hohenberger P, Labianca R, Schmoll HJ, Pahlman L, Sobrero A, Douillard JY (2007) **Endpoints in adjuvant treatment trials: a systematic review of the literature in colon cancer and proposed definitions for future trials.** J Natl Cancer Inst 99(13): 998-1003

199. Sargent DJ, Wieand HS, Haller DG, Gray R, Benedetti JK, Buyse M, Labianca R, Seitz JF, O'Callaghan CJ, Francini G, Grothey A, O'Connell M, Catalano PJ, Blanke CD, Kerr D, Green E, Wolmark N, Andre T, Goldberg RM, De GA (2005) **Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials.** J Clin Oncol 23(34): 8664-8670

200. Platell CF, Semmens JB (2004) **Review of survival curves for colorectal cancer.** Dis Colon Rectum 47(12): 2070-2075

201. Greenman C, Wooster R, Futreal PA, Stratton MR, Easton DF (2006) **Statistical analysis of pathogenicity of somatic mutations in cancer.** Genetics 173(4): 2187-2198

202. Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, Szabo S, Buckhaults P, Farrell C, Meeh P, Markowitz SD, Willis J, Dawson D, Willson JK, Gazdar AF, Hartigan J, Wu L, Liu C, Parmigiani G, Park BH, Bachman KE, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE (2006) **The consensus coding sequences of human breast and colorectal cancers.** Science 314(5797): 268-274

203. Torkamani A, Schork NJ (2008) **Prediction of cancer driver mutations in protein kinases.** Cancer Res 68(6): 1675-1682

204. Duval A, Rolland S, Compoint A, Tubacher E, Iacopetta B, Thomas G, Hamelin R (2001) **Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers.** Hum Mol Genet 10(5): 513-518

205. Duval A, Hamelin R (2002) **Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability.** Cancer Res 62(9): 2447-2454

206. Esteller M (2007) **Cancer epigenomics: DNA methylomes and histone-modification maps.** Nat Rev Genet 8(4): 286-298

207. Schuebel KE, Chen W, Cope L, Glockner SC, Suzuki H, Yi JM, Chan TA, Van NL, Van CW, van den BS, van EM, Ting AH, Jair K, Yu W, Toyota M, Imai K, Ahuja N, Herman JG, Baylin SB (2007) **Comparing the DNA hypermethylome with gene mutations in human colorectal cancer.** PLoS Genet 3(9): 1709-1723

208. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, Wright FA, Feramisco JD, Peltomaki P, Lang JC, Schuller DE, Yu L, Bloomfield CD, Caligiuri MA, Yates A, Nishikawa R, Su HH, Petrelli NJ, Zhang X, O'Dorisio MS, Held WA, Cavenee WK, Plass C (2000) **Aberrant CpG-island methylation has non-random and tumour-type-specific patterns.** Nat Genet 24(2): 132-138

209. Liu ET (2008) **Functional genomics of cancer.** Curr Opin Genet Dev 18(3): 251-256

210. Buffart TE, Overmeer RM, Steenbergen RD, Tijssen M, van Grieken NC, Snijders PJ, Grabsch HI, van d, V, Carvalho B, Meijer GA (2008) **MAL promoter hypermethylation as a novel prognostic marker in gastric cancer.** Br J Cancer 99(11): 1802-1807

Paper Ia

Guro E Lind, Terje Ahlquist, Ragnhild A Lothe.

DNA Hypermethylation of MAL: A Promising Diagnostic Biomarker for Colorectal Tumors.

Gastroenterology. 2007. Apr;132(4):1631-2

CORRESPONDENCE

DNA Hypermethylation of *MAL*: A Promising Diagnostic Biomarker for Colorectal Tumors

Dear Sir:

We found the article by Mori et al[1] recently published in GASTROENTEROLOGY of great interest. In this paper, several novel promoter methylation target genes were identified in colon cancer and among them, *MAL*, encoding a T-cell differentiation protein. This gene was also identified as a potential target in a recent study from our group using the same methodologic approach as Mori et al (microarray-based gene expression analyses before and after 5-aza-2′-deoxycytidine treatment of cell lines).[2] Only responding genes with concomitant reduced expression in in vivo tumors were selected for detailed DNA methylation analysis. We report here hypermethylation of *MAL* in an exceptionally high frequency among malignant (83%; 40 of 48 carcinomas) as well as in benign large bowel tumors (73%; 43 of 59 adenomas) as assessed by methylation-specific polymerase chain reaction (MSP) analysis (Figure 1; primers given on request).

This high methylation frequency of *MAL* is in contrast to the 6% methylation (2 of 34 samples) reported by Mori et al[1] using real-time MSP. However, they did not perform an independent validation assay for the methylation status of the gene in question, opting for caution concerning the correctness of the initial frequency. By direct bisulphite sequencing of colon cancer cell lines, we have confirmed the DNA methylation status established by MSP in the present report, and show that the majority of CpG sites were indeed methylated in the samples identified as methylated by MSP. Furthermore, the majority of normal mucosa samples were unmethylated, only 2 of 18 normal mucosa samples taken from distant sites from the primary carcinoma, and 1 of 23 normal mucosa samples from large bowels without cancer showed weak methylation (seen as a low-intensity band compared with the positive control after gel electrophoresis). The frequency found among normal samples was significantly less than in primary adenomas ($P < .0001$) and carcinomas ($P < .0001$). These results suggest that hypermethylation of *MAL* is suitable as an early diagnostic marker of primary or recurrent colorectal tumors.

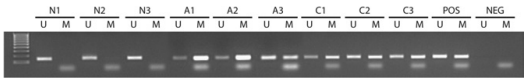


Figure 1. Representative methylation-specific polymerase chain reaction (MSP) results from the analysis of *MAL* in 3 normal mucosa samples, 3 adenomas, and 3 carcinomas. A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. *Abbreviations:* A, adenoma; C, carcinoma, N normal mucosa; POS, positive control consisting of normal blood (control for unmethylated samples) and in vitro methylated DNA (control for methylated samples); NEG, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product. The illustration is a merge of 2 gel panels as the adenomas were run on a separate gel.

Early detection of disease can result in improved clinical outcome for most types of cancer and identification of cancer-associated aberrant gene methylation represents promising novel biomarkers.[3] For colorectal cancer, initial studies have identified the presence of aberrantly methylated DNA in patient blood and feces. To our knowledge, only 2 of the genes screened for methylation in fecal DNA, *VIM* (vimentin) and *SFRP2*, have shown high sensitivity and specificity.[4,5] In general, the sensitivity and specificity of existing early markers remain suboptimal, independent of whether they are detecting DNA sequence changes or DNA modifications by methylation. Genes aberrantly hypermethylated at high frequencies already in benign tumors and only rarely in normal mucosa are good candidate diagnostic biomarkers owing to the potential clinical benefit of early detection of high-risk adenomas as well as early stages of carcinomas. Promoter hypermethylation of *MAL*, shown to be present in the vast majority of colorectal adenomas and carcinomas and only rarely in normal mucosa, therefore represents another promising early diagnostic marker that should be further studied in fecal and serum DNA samples and possibly included in a panel of biomarkers for noninvasive testing.

GURO E. LIND
TERJE AHLQUIST
RAGNHILD A. LOTHE
Department of Cancer Prevention
Institute for Cancer Research
Rikshospitalet–Radiumhospitalet Medical Centre
and Centre for Cancer Biomedicine
University of Oslo
Oslo, Norway

1. Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM, Meltzer SJ. A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer. Gastroenterology 2006;131:797–808.
2. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, Lothe RA. *ADAMTS1*, *CRABP1*, and *NR3C1* identified as epigenetically deregulated genes in colorectal tumorigenesis. Cell Oncol 2006;28:259–272.

3. Laird PW. The power and the promise of DNA methylation markers. Nat Rev Cancer 2003;3:253–266.
4. 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.
5. 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.

doi:10.1053/j.gastro.2007.03.003

Reply. We were pleased to learn that our recent report on aberrant *MAL* methylation in colon cancer was independently confirmed.[1] However, the authors appear to have made a critical oversight in interpreting both their own data and our study regarding distinct diagnostic criteria for positive methylation that determine the prevalence of "methylated" tumors.

We used quantitative real-time MSP (qMSP) assay, a sensitive, sequence-specific and quantitative assay.[2] The methylation level in a given specimen was measured as the ratio of DNA molecules demonstrating methylation at all CpG sites within primers and Taqman probe sequences relative to the total number of DNA molecules measured by a CpG-free β-actin PCR amplicon (methylation index [MI]).[1] As described in our report, we diagnosed a specimen as methylated only when this MI exceeded 0.2 in order to exclude low-level methylation events occurring in a mere minority of cells.[1] This strict criterion was chosen because our scope was to detect methylation events likely to be associated with mRNA silencing and which occur in the majority of the cells in the specimen.

The authors used qualitative MSP, a sensitive but nonquantitative method.[3] Qualitative MSP is limited in its ability to discriminate low-level from biologically relevant, high-level methylation. If we had applied a low MI threshold criterion (MI = 0.01) comparable to the standard qualitative MSP band visible on a gel, to our dataset, the prevalence of methylation in normal mucosae and primary tumors would have risen to a frequency similar to the authors' data (12% and 68%, respectively). Therefore, we agree with the authors that low-level *MAL* methylation may constitute an emerging potential early detection biomarker for colon cancer meriting further investigation. However, we conclude that low-level methylation is not likely to influence mRNA expression and is unsuitable for studies emphasizing biologically relevant methylation. In support of our conclusion, reexpression of *MAL* mRNA was observed in 5-aza-2′-deoxycitidine (Aza-C)-treated HT29 cells, although these cells still maintained low-level *MAL* methylation after Aza-C treatment (MI = 0.09).[1]

Finally, we consider validation with bisulfite sequencing to have been unnecessary in our qMSP study. qMSP ensures highly efficient discrimination specific from nonspecific PCR products because of methylated sequence-specific detection using TaqMan probes containing CpGs.[2] As quality controls, we confirmed that our *MAL*-qMSP-amplicon did not amplify from unconverted DNA or unmethylated DNA. Furthermore, we observed proper and sensitive amplification from fully methylated positive control DNA, even after 625-fold dilution.

In conclusion, both our study and the authors' own data describe essentially the same phenomenon, but with vastly different diagnostic criteria. We would like to emphasize the importance of appropriately designing and interpreting epigenetic studies of cancers, in light of the distinct nature of each analytical method and the focus of each study.

YURIKO MORI
FUMIAKI SATO
STEPHEN J. MELTZER
Department of Medicine
Johns Hopkins School of Medicine
Baltimore, Maryland

1. Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM, Meltzer SJ. A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer. Gastroenterology 2006;131:797–808.
2. Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, Danenberg PV, Laird PW. MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 2000;28:E32.
3. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–9826.

doi:10.1053/j.gastro.2007.03.004

AGA Position Statement of Computed Tomographic Colonography

Dear Sir:

We read with interest the recent American Gastroenterological Association (AGA) Institute's statement entitled "Position of the American Gastroenterological Association (AGA) Institute on Computed Tomographic Colonography."[1] This statement announces that "the AGA Institute has convened a task force to develop training standards for gastroenterologists' performance of CT colonography." As a combined gastroenterologist (PL) and radiologist (AF) team, we believe that this initiative is seriously misguided.

Interpretation of computed tomographic colonography (CTC), as currently practiced, requires the interpreting physician to interact on a workstation with a volumetric CT dataset that usually consists of images acquired in both supine and prone positions. At a minimum the colon is examined with a volume-rendered 3D endoluminal display, synchronized with a simultaneous multiplanar 2D display (primary 3D read), and/or a 2D simultaneous axial supine/prone display with 3D of points of interest (primary 2D read).

Paper Ib

Guro E Lind, Terje Ahlquist, Matthias Kolberg, Marianne Berg, Mette Eknæs, Miguel A Alonso, Anne Kallioniemi, Gunn I Meling, Rolf I Skotheim, Torleiv O Rognum, Espen Thiis-Evensen and Ragnhild A Lothe.

Hypermethylated MAL gene – a silent marker of early colon tumorigenesis.

Journal of Translational Medicine. 2008 Mar 17;6:13.

Journal of Translational Medicine BioMed Central

Research

Open Access

Hypermethylated *MAL* gene – a silent marker of early colon tumorigenesis

Guro E Lind[1,2], Terje Ahlquist[1,2], Matthias Kolberg[1,2], Marianne Berg[1,2], Mette Eknæs[1,2], Miguel A Alonso[3], Anne Kallioniemi[4], Gunn I Meling[5,6], Rolf I Skotheim[1,2], Torleiv O Rognum[7], Espen Thiis-Evensen[8] and Ragnhild A Lothe*[1,2]

Address: [1]Department of Cancer Prevention, Institute for Cancer Research, The Norwegian Radium Hospital, Rikshospitalet University Hospital, Oslo, Norway, [2]Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway, [3]Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas y Universidad Autónoma de Madrid, Madrid, Spain, [4]Laboratory of Cancer Genetics, Institute of Medical Technology, Tampere University Hospital and University of Tampere, Tampere, Finland, [5]Department of Urology, Akershus University Hospital, Lørenskog, Norway, [6]Faculty of Medicine, University of Oslo, Oslo, Norway, [7]Institute of Forensic Medicine, Rikshospitalet University Hospital, Oslo, Norway and [8]Medical Department, Rikshospitalet University Hospital, Oslo, Norway

Email: Guro E Lind - Guro.Elisabeth.Lind@rr-research.no; Terje Ahlquist - terje.c.ahlquist@rr-research.no; Matthias Kolberg - matthias.kolberg@rr-research.no; Marianne Berg - marianne.berg@rr-research.no; Mette Eknæs - metteek@rr-research.no; Miguel A Alonso - maalonso@cbm.uam.es; Anne Kallioniemi - anne.kallioniemi@uta.fi; Gunn I Meling - gi@meling.net; Rolf I Skotheim - rolf.i.skotheim@rr-research.no; Torleiv O Rognum - t.o.rognum@medisin.uio.no; Espen Thiis-Evensen - Espen.Thiis-Evensen@rikshospitalet.no; Ragnhild A Lothe* - rlothe@rr-research.no

* Corresponding author

Published: 17 March 2008

Journal of Translational Medicine 2008, **6**:13 doi:10.1186/1479-5876-6-13

Received: 8 January 2008
Accepted: 17 March 2008

This article is available from: http://www.translational-medicine.com/content/6/1/13

Abstract

Background: Tumor-derived aberrantly methylated DNA might serve as diagnostic biomarkers for cancer, but so far, few such markers have been identified. The aim of the present study was to investigate the potential of the *MAL* (T-cell differentiation protein) gene as an early epigenetic diagnostic marker for colorectal tumors.

Methods: Using methylation-specific polymerase chain reaction (MSP) the promoter methylation status of *MAL* was analyzed in 218 samples, including normal mucosa (n = 44), colorectal adenomas (n = 63), carcinomas (n = 65), and various cancer cell lines (n = 46). Direct bisulphite sequencing was performed to confirm the MSP results. *MAL* gene expression was investigated with real time quantitative analyses before and after epigenetic drug treatment. Immunohistochemical analysis of MAL was done using normal colon mucosa samples (n = 5) and a tissue microarray with 292 colorectal tumors.

Results: Bisulphite sequencing revealed that the methylation was unequally distributed within the *MAL* promoter and by MSP analysis a region close to the transcription start point was shown to be hypermethylated in the majority of colorectal carcinomas (49/61, 80%) as well as in adenomas (45/63, 71%). In contrast, only a minority of the normal mucosa samples displayed hypermethylation (1/23, 4%). The hypermethylation of *MAL* was significantly associated with reduced or lost gene expression in *in vitro* models. Furthermore, removal of the methylation re-induced gene expression in colon cancer cell lines. Finally, MAL protein was expressed in epithelial cells of normal colon mucosa, but not in the malignant cells of the same type.

Conclusion: Promoter hypermethylation of *MAL* was present in the vast majority of benign and malignant colorectal tumors, and only rarely in normal mucosa, which makes it suitable as a diagnostic marker for early colorectal tumorigenesis.

Background

Epigenetic changes – non-sequence-based alterations that are inherited through cell division [1] – are frequently seen in human cancers, and likewise as genetic alterations they may lead to disruption of gene function. In colorectal cancer, several tumour suppressor genes have been identified to be epigenetically inactivated by CpG island promoter hypermethylation, including the DNA mismatch repair gene *MLH1* [2-4], the gatekeeper *APC* [5], and the cell cycle inhibitor *CDKN2A* [6], to mention some. In addition to contributing to, or accompanying, the stepwise development of malignant colorectal carcinomas from benign adenomas, aberrant DNA methylation holds great promises for cancer diagnostics [7]. Based on the ubiquity of aberrant promoter methylation and the ability to detect this methylated DNA in body fluids, such as blood, the presence of this altered DNA may represent potential diagnostic biomarkers for cancer. For non-invasive detection of colorectal tumours, stool is the obvious source of DNA for such investigations and several studies have identified cancer-derived aberrant DNA hypermethylation using this approach [8-10]. However, the sensitivity and specificity of these tests are still suboptimal and would benefit from incorporating additional biomarkers.

We recently published a list of promising novel target genes for hypermethylation in colorectal tumours [11]. Among these was the *MAL* (T-cell differentiation protein) gene, and we then communicated that the CpG rich promoter of *MAL* seemed to be hypermethylated in the majority of colorectal tumours [12].

The *MAL* gene, which was initially isolated and cloned in 1987, maps to chromosome band 2cen-q13, encodes a 17 kDa integral membrane protein, and contains a CpG island [13,14]. Originally, expression of *MAL* was found in intermediate and late stages of T-cell differentiation, and MAL was suggested to play a role in membrane signalling [15]. In recent years, MAL has also been shown to play a role in apical transport, which is a polarized transport of lipids and proteins to the apical (external facing) membrane in certain cell types [16]. Such polarized transport is essential for the proper functioning of epithelial cells, and the neoplastic transformation process is frequently associated with loss of this polarized phenotype [16]. Finally, MAL has been shown to possess tumour suppressor capabilities by suppressing motility, invasion, and tumorigenicity and enhance apoptosis in oesophageal cancer [17].

In the present study, we have compared the promoter methylation status of *MAL* in a large series of normal colorectal mucosa samples, with those of benign and malignant colorectal tumours. Furthermore, RNA and/or protein expression levels of MAL were determined in *in vivo* tumours as well as in *in vitro* models, the latter also including various cancer types. The findings were used to decide whether or not methylated *MAL* is suitable as a diagnostic marker for early colorectal tumorigenesis.

Methods

Patients and cell lines

DNA from 218 fresh-frozen samples was subjected to methylation analysis, including 65 colorectal carcinomas (36 micro satellite stable; MSS, and 29 with micro satellite instability; MSI) from 64 patients, 63 adenomas, median size 8 mm, range 5–50 mm (61 MSS and 2 MSI) from 52 patients, 21 normal mucosa samples from 21 colorectal cancer patients (taken from distant sites from the primary carcinoma), and another 23 normal colorectal mucosa samples from 22 cancer-free individuals, along with 20 colon cancer cell lines (11 MSS and 9 MSI), and 26 cancer cell lines from various tissues (breast, kidney, ovary, pancreas, prostate, and uterus; Table 1). The mean age at diagnosis was 70 years (range 33 to 92) for patients with carcinoma, 67 years (range 62 to 72) for persons with adenomas, 64 years (ranging from 24 to 89) for the first group of normal mucosa donors, and 54 years (ranging from 33 to 86) for the second group of normal mucosa donors. The colorectal carcinomas and normal samples from cancer patients were obtained from an unselected prospective series collected from seven hospitals located in the South-East region of Norway [18]. The adenomas were obtained from individuals attending a population based sigmoidoscopic screening program for colorectal cancer [19]. The normal mucosa samples from cancer-free individuals were obtained from deceased persons, and the majority of the total set of normal samples (27/44) consisted of mucosa only, whereas the remaining samples were taken from the bowel wall. Additional clinico-pathological data for the current tumour series include gender and tumour location, as well as polyp size and total number of polyps per individual for the adenoma series.

All samples were retrieved from approved research biobanks and are part of research projects approved according to national guidelines (Biobank; registered at the Norwegian Institute of Public Health. Projects: Regional Ethics Committee and National Data Inspectorate).

Two colon cancer cell lines, HCT15 and HT29, were subjected to treatment with the demethylating drug 5-aza-2'deoxycytidine (1 μM for 72 h), the histone deactetylase inhibitor trichostatin A (0.5 μM for 12 h) and a combination of both (1 μM 5-aza-2'deoxycytidine for 72 h, 0.5 μM trichostatin A added the last 12 h).

Table 1: Promoter methylation status of *MAL* in cell lines of various tissues.

Cell line	Tissue	Promoter methylation status	Methylation frequency
BT-20	Breast	M	57%
BT-474	Breast	U/M	
Hs 578T	Breast	U	
SK-BR-3	Breast	U	
T-47D	Breast	U/M	
ZR-75-1	Breast	U	
ZR-75-38	Breast	M	
Co115	Colon	M	95%
HCT15	Colon	M	
HCT116	Colon	M	
LoVo	Colon	M	
LS174T	Colon	M	
RKO	Colon	M	
SW48	Colon	M	
TC7	Colon	M	
TC71	Colon	M	
ALA	Colon	M	
Colo320	Colon	M	
EB	Colon	M	
FRI	Colon	U/M	
HT29	Colon	M	
IS1	Colon	M	
IS2	Colon	M	
IS3	Colon	M	
LS1034	Colon	M	
SW480	Colon	M	
V9P	Colon	U	
ACHN	Kidney	U	50%
Caki-1	Kidney	U	
Caki-2	Kidney	M	
786-O	Kidney	U/M	
ES-2	Ovary	U/M	50%
OV-90	Ovary	U/M	
Ovcar-3	Ovary	U	
SK-OV-3	Ovary	U	
AsPC-1	Pancreas	M	67%
BxPC-3	Pancreas	U	
CFPAC-1	Pancreas	U	
HPAF-II	Pancreas	M	
PaCa-2	Pancreas	M	
Panc-1	Pancreas	U/M	
LNCaP	Prostate	U	0%
AN3 CA	Uterus	U/M	75%
HEC-1-A	Uterus	M	
KLE	Uterus	U	
RL95-2	Uterus	M	

The promoter methylation status of the individual cell lines was assessed by methylation-specific polymerase chain reaction (MSP). The methylation frequency reflects the number of methylated (M and U/M) samples from each tissue. Abbreviations: U, unmethylated; M, methylated.

Bisulphite treatment and methylation-specific polymerase chain reaction (MSP)

DNA from primary tumours and normal mucosa samples was bisulphite treated as previously described [11,20], whereas DNA from colon cancer cell lines was bisulphite treated using the EpiTect bisulphite kit (Qiagen Inc., Valencia, CA, USA). The promoter methylation status of *MAL* was analyzed by methylation-specific polymerase chain reaction (MSP) [21], using the HotStarTaq DNA polymerase (Qiagen). All results were confirmed with a second independent round of MSP. Human placental DNA (Sigma Chemical Co, St. Louis, MO, USA) treated *in vitro* with *Sss1* methyltransferase (New England Biolabs Inc., Beverly, MA, USA) was used as a positive control for the methylated MSP reaction, whereas DNA from normal lymphocytes was used as a positive control for unmethylated alleles. Water was used as a negative control in both reactions. The primers were designed with MethPrimer [22] and their sequences are listed in Table 2, along with the product fragment lengths and primer locations.

Bisulphite sequencing

All colon cancer cell lines (n = 20) were subjected to direct bisulphite sequencing of the *MAL* promoter [23]. Two fragments were amplified: fragment A, covering bases -68 to 168 relative to the transcription start point (overlapping with our MSP product), and fragment B covering bases -427 to -23. Fragment A covered altogether 24 CpG sites and was amplified using the HotStarTaq DNA polymerase and 35 PCR cycles. Fragment B covered altogether 32 CpG sites and was amplified using the same polymerase and 36 PCR cycles. The primer sequences are listed in Table 2. Excess primer and nucleotides were removed by ExoSAP-IT treatment following the protocol of the manufacturer (GE Healthcare, USB Corporation, Ohio, USA). The purified products were subsequently sequenced using the dGTP BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA) in an AB Prism 3730 sequencer (Applied Biosystems). The approximate amount of methyl cytosine of each CpG site was calculated by comparing the peak height of the cytosine signal with the sum of the cytosine and thymine peak height signals, as previously described [24]. CpG sites with ratios ranging from 0 – 0.20 were classified as unmethylated, CpG sites within the range 0.21 – 0.80 were classified as partially methylated, and CpG sites ranging from 0.81 – 1.0 were classified as hypermethylated.

cDNA preparation and real-time quantitative gene expression

Total RNA was extracted from cell lines (n = 46), tumours (n = 16), and normal tissue (n = 3) using Trizol (Invitrogen, Carlsbad, CA, USA) and the RNA concentration was determined using ND-1000 Nanodrop (NanoDrop Technologies, Wilmington, DE, USA). For each sample, total RNA was converted to cDNA using a High-Capacity cDNA Archive kit (Applied Biosystems), including random primers. *MAL* (Hs00242749_m1 and Hs00360838_m1) and the endogenous controls *ACTB* (Hs99999903_m1) and *GUSB* (Hs99999908_m1) were amplified separately in 96 well fast plates following the recommended protocol (Applied Biosystems), and the real time quantitative gene expression was measured by the 7900 HT Sequence Detection System (Applied Biosystems). All samples were analyzed in triplicate, and the median value was used for data analysis. The human universal reference RNA (containing a mixture of RNA from ten different cell lines; Stratagene) was used to generate a standard curve, and the resulting quantitative expression levels of *MAL* were normalized against the mean value of the two endogenous controls.

Tissue microarray

For *in situ* detection of protein expression in colorectal cancers, a tissue microarray (TMA) was constructed, based on the technology previously described [25]. Embedded in the TMA are 292 cylindrical tissue cores (0.6 mm in diameter) from ethanol-fixed and paraffin embedded tumour samples derived from 281 individuals. Samples from the same patient series has been examined for various biological variables and clinical end-points [18,26-28]. In addition, the array contains normal tissues from kidney, liver, spleen, and heart as controls. Ethanol-fixed

Table 2: PCR primers used for MSP and bisulphite sequencing.

Primer set	Sense primer	Antisense primer	Frg. Size, bp	An. Temp	Fragment location*
MAL MSP-M	TTCGGGTTTTTTTGTTTTTAATTC	GAAAACCATAACGACGTACTAACGT	139	56	-71 to 68
MAL MSP-U	TTTTGGGTTTTTTTGTTTTTAATTT	ACAAAAACCATAACAACATACTAACATC	142	56	-72 to 70
MAL BS_A	GGGTTTTTTTGTTTTTAATT	ACCAAAAACCACTCACAAACTC	236	53	-68 to 168
MAL BS_B	GGAAAAATGAAGGAGATTTAAATTT	AATAACCTAAACRCCCCC	404	50	-427 to -23

Abbreviations: MSP, methylation-specific polymerase chain reaction; BS, bisulphite sequencing; M, methylated-specific primers; U, unmethylated-specific primers; Frg. Size, fragment size; An. Temp, annealing temperature (in degrees celsius). *Fragment location lists the start and end point (in base pairs) of each fragment relative to the transcription start point provided by NCBI (RefSeq ID NM_002371).

normal colon tissues from four persons with no known history of colorectal cancer were obtained separately.

Immunohistochemical in situ *protein expression analysis*

Five μm thick sections of the TMA blocks were transferred onto glass slides for immunohistochemical analyses. The sections were deparaffinized in a xylene bath for 10 minutes and rehydrated via a series of graded ethanol baths. Heat-induced epitope retrieval was performed by heating in a microwave oven at full effect (850 W) for 5 minutes followed by 15 minutes at 100 W immersed in 10 mM citrate buffer at pH 6.0 containing 0.05% Tween-20. After cooling to room temperature, the immunohistochemical staining was performed according to the protocol of the DAKO Envision+™ K5007 kit (Dako, Glostrup, Denmark). The primary antibody, mouse clone 6D9 anti-MAL [29], was used at a dilution of 1:5000, which allowed for staining of kidney tubuli as positive control, while the heart muscle tissue remained unstained as negative control [30]. The slides were counterstained with haematoxylin for 2 minutes and then dehydrated in increasing grades of ethanol and finally in xylene. Results from the immunohistochemistry were obtained by independent scoring by one of the authors and a reference pathologist.

Statistics

All *P* values were derived from two tailed statistical tests using the SPSS 13.0 software (SPSS, Chicago, IL, USA). Fisher's exact test was used to analyze 2 × 2 contingency tables. A 2 × 3 table and Chi-square test was used to analyze the potential association between quantitative gene expression of *MAL* and promoter methylation status. Samples were divided into two categories according to their gene expression levels: low expression included samples with gene expression equal to, or lower than, the median value across all cell lines or all tumours, high expression included samples with gene expression higher that the median. The methylation status was divided into three categories: unmethylated, partial methylation, and hypermethylated.

Results

Promoter methylation status of MAL *in tissues and cell lines*

The promoter methylation status of *MAL* was analyzed with MSP (Figure 1). One of 23 (4%) normal mucosa samples from non-cancerous donors and two of 21 (10%) normal mucosa samples taken in distance from the primary tumour were methylated but displayed only low-intensity band compared with the positive control after gel electrophoresis. Forty-five of 63 (71%) adenomas and 49/61 (80%) carcinomas showed promoter hypermethylation. Nineteen of twenty colon cancer cell lines (95%), and 15/26 (58%) cancer cell lines from various tissues (breast, kidney, ovary, pancreas, prostate, and uterus) were hypermethylated (Table 1 lists tissue-specific frequencies).

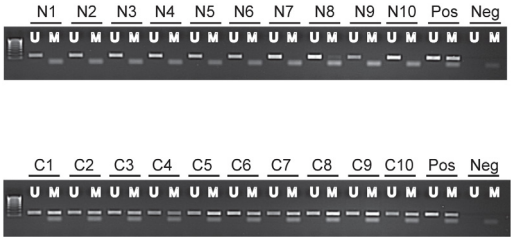


Figure 1
Methylation status of the *MAL* promoter in normal colon mucosa samples and colorectal carcinomas. Representative results from methylation-specific polymerase chain reaction are shown. A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. N, normal mucosa; C, carcinoma; Pos, positive control (unmethylated reaction: DNA from normal blood, methylated reaction: *in vitro* methylated DNA); Neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product.

The hypermethylation frequency found in normal samples was significantly lower than in adenomas ($P < 0.0001$) and carcinomas ($P < 0.0001$). Hypermethylation of the *MAL* promoter was not associated with MSI status, gender, or age in neither malignant nor benign tumours. Among carcinomas, tumours with distal location in the bowel (left side and rectum) were more frequently hypermethylated than were tumours with proximal location, although not statistically significant ($P = 0.088$). Among adenomas, no significant association could be found between promoter methylation status of *MAL* and polyp size or number.

Bisulphite sequencing verification of the promoter methylation status of MAL

Two overlapping fragments of the *MAL* promoter were bisulphite sequenced in 20 colon cancer cell lines. The results are summarized in Figure 2, and representative raw data can be seen in Figure 3. A good association was seen between the methylation status, as assessed by MSP, and the bisulphite sequences of the overlapping fragment A. However, in fragment B there was poor association with the MSP data. For this fragment, which is located farther upstream relative to the transcription start point, several consecutive CpG sites were frequently unmethylated and/or partially methylated. This held true also in cell lines shown to be heavily methylated around the transcription start point (fragment A; Figure 2).

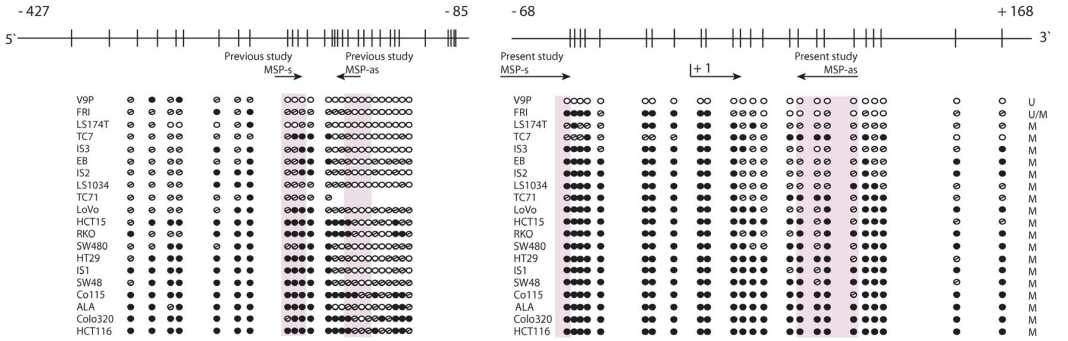


Figure 2

Site specific methylation within the *MAL* promoter. Bisulphite sequencing of the *MAL* promoter verifies methylation status assessed by methylation-specific polymerase chain reaction. The upper part of the figure is a schematic presentation of the CpG sites successfully amplified by the two analyzed bisulphite sequencing fragments, A (-68 to +168; to the right) and B (-427 to -85; to the left). The transcription start site is represented by +1 and the vertical bars indicate the location of individual CpG sites. The two arrows indicate the location of the MSP primers in the present study and a previously published study analyzing promoter methylation of *MAL* [31]. For the lower part of the figure, filled circles represent methylated CpGs; open circles represent unmethylated CpGs; and open circles with a slash represent partially methylated sites (the presence of approximately 20–80% cytosine, in addition to thymine). The column of U, M and U/M at the right side of this lower part lists the methylation status of the respective cell lines as assessed by us using MSP analyses. Abbreviations: MSP, methylation-specific PCR; s, sense; as, antisense; U, unmethylated; M, methylated; U/M, presence of both unmethylated and methylated band.

Real-time quantitative gene expression

The level of *MAL* mRNA expression in cell lines (n = 46), primary colorectal carcinomas (n = 16), and normal mucosa (n = 3) was assessed by quantitative real time PCR. There was a strong association between *MAL* promoter hypermethylation and reduced or lost gene expression among cell lines ($P = 0.041$; Figure 4). Furthermore, the gene expression of *MAL* was up-regulated in colon cancer cell lines after promoter demethylation induced by the combined treatment 5-aza-2'-deoxycytidine and trichostatin A (Figure 5). Treatment with the deacetylase inhibitor trichostatin A alone did not increase *MAL* expression, whereas treatment with the DNA demethylating 5-aza-2'-deoxycytidine led to high expression in HT29 cells, but more moderate levels in HCT15 cells (Figure 5). Among primary colorectal carcinomas, those harbouring promoter hypermethylation of *MAL* (n = 13) expressed somewhat lower levels of *MAL* mRNA compared with the unmethylated tumours (n = 3), although not statistically significant (Figure 4).

MAL protein expression is lost in colorectal carcinomas

To evaluate the immunohistochemistry analyses of MAL, kidney and heart muscle tissues were included as positive and negative controls, respectively (Figure 6A–B) [30]. From the 231 scorable colorectal tissue cores, *i.e.* those containing malignant colorectal epithelial tissue, 198 were negative for MAL staining (Figure 6C–D). Twenty-nine of these had positive staining in non-epithelial tissue components within the same tissue cores, mainly in neurons and blood vessels (not shown). In comparison, all the sections of normal colon tissue contained positive staining for MAL in the epithelial cells (Figure 6E–F).

Discussion

In the present study, we have demonstrated that a sequence within the *MAL* promoter close to the transcription start is hypermethylated in the vast majority of malignant, as well as in benign colorectal tumours, in contrast to normal colon mucosa samples which are unmethylated, and we contend that *MAL* remains a promising diagnostic biomarker for early colorectal tumorigenesis [12]. The adenomas and carcinomas analyzed in the present study are from unselected clinical series and are therefore representative for the average risk population. However, the equal distribution between MSI and MSS carcinomas in the present study is not representative for a consecutive series.

Hypermethylation of *MAL* has, by quantitative methylation-specific polymerase chain reaction (MSP), previously been shown by others to be present only in a small fraction (6%, 2/34) of colon carcinomas [31], even though the expression of *MAL* was reported to be reduced/lost in the majority of colorectal tumours [11,17,31]. In contrast, we report here a significantly higher methylation fre-

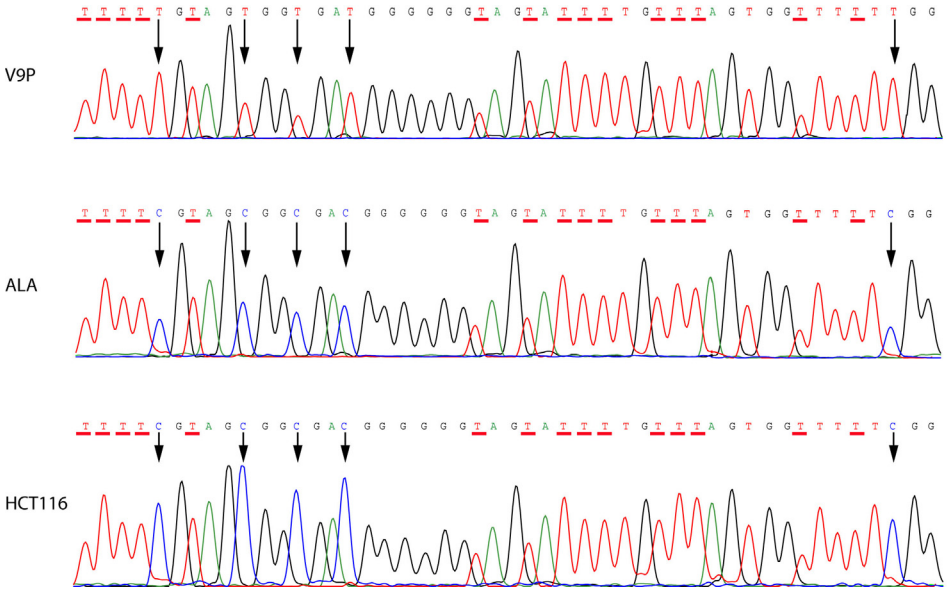


Figure 3
The "bisulphite sequence" of the *MAL* promoter. Representative bisulphite sequencing electropherograms of the *MAL* promoter in colon cancer cell lines. A subsection of the bisulphite sequence electropherogram, covering CpG sites +11 to +15 relative to transcription start. Cytosines in CpG sites are indicated by a black arrow, whereas cytosines that have been converted to thymines are underlined in red. The *MAL* promoter sequencing electropherograms illustrated here, are from the unmethylated V9P cell line and the hypermethylated ALA and HCT116.

quency of *MAL* in both benign and malignant colorectal tumours (71% in adenomas and 80% in carcinomas). The discrepancy in methylation frequencies between the present report and the previous study by Mori and co-workers [31] is probably a consequence of study design. From direct bisulphite sequencing of colon cancer cell lines, we have now shown that the DNA methylation of *MAL* is unequally distributed within the CpG island of its promoter (Figure 2). CpG islands often span more than one kilobase of the gene promoter, and the methylation status within this region is sometimes mistakenly assumed to be equally distributed. This is exemplified by the *MLH1* gene in which hypermethylation of a limited number of CpG sites approximately 200 base pairs upstream of the transcription start point invariably correlates with the lack of gene expression, while other sites do not [32,33]. Since the results of an MSP analysis rely on the match or mismatch of the unmethylated and methylated primer sequences to bisulphite treated DNA, one should ensure that the primers anneal to relevant CpG sites in the gene promoter. In the present study, we designed the MSP primers close to the transcription start point of the gene (-72 to +70) and found, by bisulphite sequencing, concordance between the overall methylation status of *MAL* as assessed by MSP and the methylation status of the individual CpG sites covered by our MSP primer set (Figure 2). This part of the CpG island was hypermethylated in the majority of colon cancer cell lines (95%). We also found that these cell lines, as well as those of other tissues, showed loss of *MAL* RNA expression from quantitative real time analyses, and that removal of DNA hypermethylation by the combined treatment of 5-aza-2'-deoxycytidine and Trichostatin A re-induced the expression of *MAL* in colon cancer cell lines (Figure 5). Furthermore, by analyzing a large series of clinically representative samples by protein immunohistochemistry we confirmed that the expression of MAL was lost in malignant colorectal epithelial cells as compared to normal mucosa.

We have further analyzed the same region of the *MAL* promoter as Mori *et al.*, which is located -206 to -126 base pairs upstream of the transcription start point [31]. By direct bisulphite sequencing, we showed that only a minority of the CpG sites covered by the Mori antisense primer were methylated in the 19 colon cancer cell lines

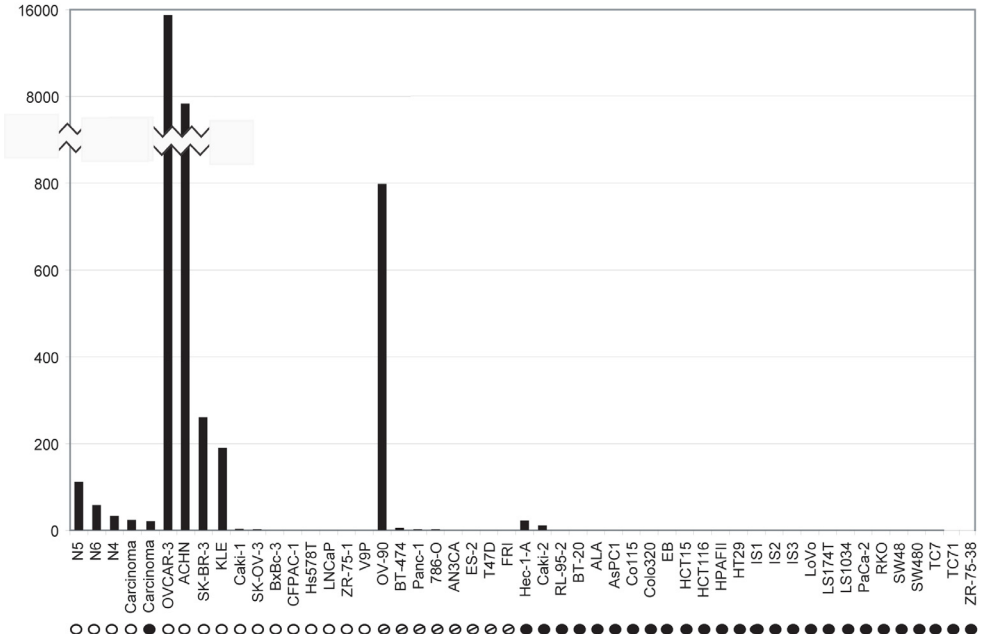


Figure 4

***MAL* expression in cancer cell lines and colorectal carcinomas**. Promoter hypermethylation of *MAL* was associated with reduced or lost gene expression in *in vitro* models. The quantitative gene expression level of *MAL* is displayed as a ratio between the average of two *MAL* assays (detecting various splice variants) and the average of the two endogenous controls, *GUSB* and *ACTB*. The value has been multiplied by a factor of 1000. Below each sample the respective methylation status is shown, as assessed by methylation-specific polymerase chain reaction. Filled circles represent promoter hypermethylation of *MAL*, open circles represent unmethylated *MAL*, and open circles with a slash represent the presence of both unmethylated and methylated alleles. Colorectal carcinomas are divided in an unmethylated group (n = 3) and a hypermethylated group (n = 13), and the median expression is displayed here. The tissue of origin for the individual cell lines can be found in table 1.

that were heavily methylated around the transcription start point (Figure 2). We therefore conclude that the very low (six percent) methylation frequency initially reported for *MAL* in colon carcinomas [31] is most likely a consequence of the primer design and choice of CpG sites to be examined.

Inactivating hypermethylation of the *MAL* promoter might be prevalent also in other cancer types where low expression of *MAL* has been shown not to correlate with allelic loss or somatic mutations in the *MAL* gene [34]. In the present study, hypermethylated *MAL* was found in cancer cell lines from breast, kidney, ovary, and uterus.

The present analyses of cancer cell lines from seven tissues indicate that the hypermethylation of a limited area in the proximity of the transcription start point of *MAL* is associated with reduced or lost gene expression. However, among colorectal carcinomas and cell lines, MAL protein and gene expression seemed to be lost or reduced in all samples, including the minority with unmethylated *MAL* promoters. This underlines that loss of the MAL protein might have an important function in colorectal tumorigenesis and we hypothesize that early during colorectal neoplasia the gene is turned off by epigenetic mechanisms other than DNA methylation. The DNA methylation is subsequently recruited to the MAL promoter to "seal" the unexpressed state. Hence, it needs to be established

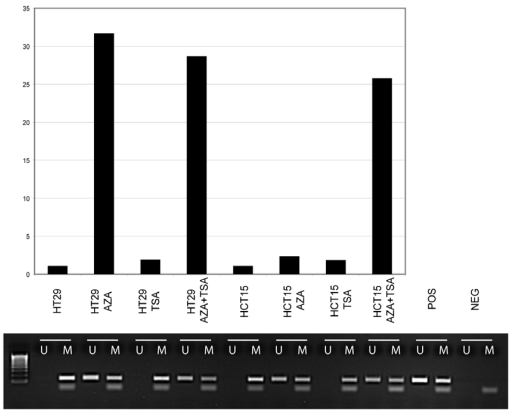


Figure 5
Up-regulation of *MAL* expression after drug treatment. Decreased promoter methylation of *MAL* followed by up-regulated mRNA expression in colon cancer cell lines was found after treatment with the demethylating 5-aza-2'deoxycytidine, alone and in combination with the deacetylase inhibitor trichostatin A. Upper panel demonstrate the relative expression values of *MAL* (linear scale) in two colon cancer cell lines, HT29 and HCT15, treated with 5-aza-2'deoxycytidine alone, trichostatin A alone, and the two drugs in combination. Lower panel illustrate *MAL* MSP results for the same samples. A visible PCR product in lanes U indicates the presence of unmethylated alleles whereas a PCR product in lanes M indicates the presence of methylated alleles. Abbreviation: AZA, 5-aza-2'deoxycytidine; TSA, trichostatin A; Pos, positive control (unmethylated reaction: DNA from normal blood, methylated reaction: *in vitro* methylated DNA); Neg, negative control (containing water as template); U, lane for unmethylated MSP product; M, lane for methylated MSP product.

whether *MAL* promoter hypermethylation is a cause or a consequence of the observed loss of gene expression in colorectal tumors. This is interesting from a biological – but not necessarily a diagnostic – perspective. The distinction between the two is supported by the fact that one of the most promising diagnostic biomarkers for colorectal cancer reported so far, DNA hypermethylation of the vimentin (*VIM*) gene, is not expected to alter the gene expression, nor to confer a selective advantage upon cancer cells in the colon, considering the lack of *VIM* expression by normal colonic epithelial cells [9].

A sensitive non-invasive screening approach for colorectal cancer could markedly improve the clinical outcome for the patient. Such a diagnostic test could in principle measure the status of a single biomarker, although multiple markers are probably needed to achieve sufficient sensitivity and specificity. Several studies have successfully detected such tumour-specific products in the faeces, and most experience has been with mutant genetic markers, including *APC*, *KRAS*, *TP53*, and BAT-26 [35]. However, one of the most promising faecal DNA tests so far consisted of a combination of a genetic DNA integrity assay and an epigenetic *VIM* methylation assay, resulting in 88% sensitivity and 82% specificity [36]. This panel might be further improved by implementing *MAL* and/or, as suggested by others, the *SFRP2* marker, which has an independent sensitivity and specificity of 77% in faecal DNA [8].

Hypermethylation of the *MAL* promoter represents, to the best of our knowledge, the most frequently hypermethylated gene among pre-malignant colorectal lesions, accompanied by low methylation frequencies in normal colon mucosa. The presence of such epigenetic changes in pre-malignant tissues might also have implications for cancer chemoprevention. By inhibiting or reversing these epigenetic alterations, the progression to a malignant phenotype might be prevented [37]. However, for the purpose of cancer risk assessment, *MAL* methylation status should be used in combination with other markers to recognize high risk adenomas.

Conclusion

Promoter hypermethylation of *MAL* remains one of the most promising diagnostic biomarkers for early detection of colorectal tumours, and, together with other biomarkers, it merits further investigation with the purpose of developing a diagnostic marker panel with the necessary sensitivity and specificity to discover colorectal neoplasia and perform a risk assessment.

Abbreviations

MAL, T-cell differentiation protein gene; MSI, micro satellite instability; MSP, methylation-specific polymerase chain reaction; MSS, micro satellite stable micro satellite instability; TMA, tissue microarray.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

GEL and TA carried out the MSP analyses and interpreted the results independent of each other. GEL additionally designed the study, carried out the bisulphite sequencing and the quantitative real-time PCR analyses, performed the statistics, and drafted the manuscript. MK carried out the immunohistochemistry analyses, interpreted the results together with an expert pathologist and contributed in manuscript preparations. MB isolated DNA from cancer cell lines. ME cultured all cell lines and treated

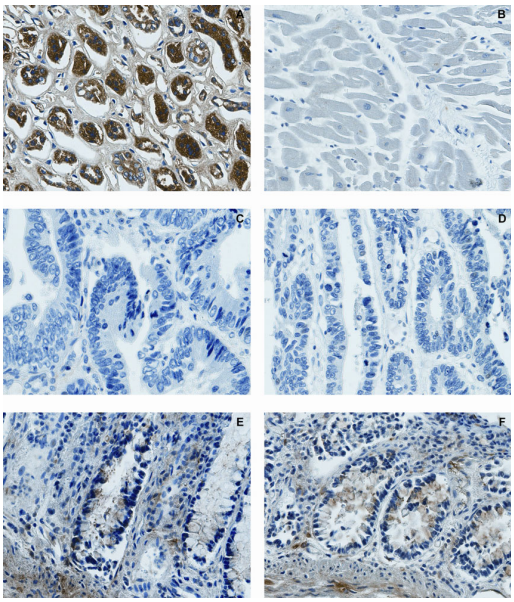


Figure 6
Lack of MAL protein expression in colorectal carcinomas. Positive cytoplasmic staining of MAL was found in kidney tubuli (A), and no staining was observed in heart muscle (B), in agreement with earlier reports [30]. The epithelial cells of colorectal carcinomas were MAL negative (C, D), whereas in normal colon tissue, cytoplasmic expression of MAL was found in both epithelia and connective tissue (E, F). All images were captured using the 40× lens (400× magnification).

colon cancer cell lines with epigenetic drugs. MAA provided the antibody for immunohistochemical analyses and contributed with scientific discussion. AK provided several of the cancer cell lines. GIM and TOR have collected the series of human primary carcinomas and normal mucosa tissue and provided all clinical and pathological information regarding these samples. RIS generated the tissue microarray and contributed in manuscript preparations. ETE has provided the series of adenoma samples and provided all clinical and pathological information regarding these samples and contributed in manuscript preparations. RAL conceived the study, participated in its design, contributed in evaluation of results, scientific discussion and in manuscript preparation. All authors have read and approved the final manuscript.

Acknowledgements

We are grateful to Vera M. Abeler for evaluating the immunohistochemical staining of the TMA tissue cores and to Stine Aske Danielsen for technical assistance. GEL is a post doc and MB is a PhD student supported by the Norwegian Cancer Society (grant A95068, RAL). The study is also supported by grants from the Norwegian Research Council (163962/V50, RAL and 161448/V40, RAL). MAA is supported by grants from the Ministerio de Educación y Ciencia (BFU2006-01925; GEN2003-20662-C07-02).

References

1. Feinberg AP, Ohlsson R, Henikoff S: **The epigenetic progenitor origin of human cancer.** *Nat Rev Genet* 2006, **7:**21-33.
2. 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.
3. 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.
4. 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.
5. Esteller M, Sparks A, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Gonzalez S, Tarafa G, Sidransky D, Meltzer SJ, Baylin SB, Herman JG: **Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer.** *Cancer Res* 2000, **60:**4366-4371.
6. Burri N, Shaw P, Bouzourene H, Sordat I, Sordat B, Gillet M, Schorderet D, Bosman FT, Chaubert P: **Methylation silencing and mutations of the p14ARF and p16INK4a genes in colon cancer.** *Lab Invest* 2001, **81:**217-229.
7. Laird PW: **The power and the promise of DNA methylation markers.** *Nat Rev Cancer* 2003, **3:**253-266.
8. Müller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, Mühlthaler M, Ofner D, Margreiter R, Widschwendter M: **Methylation changes in faecal DNA: a marker for colorectal cancer screening?** *Lancet* 2004, **363:**1283-1285.
9. 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.
10. Zhang W, Bauer M, Croner RS, Pelz JO, Lodygin D, Hermeking H, Sturzl M, Hohenberger W, Matzel KE: **DNA Stool Test for Colorectal Cancer: Hypermethylation of the Secreted Frizzled-Related Protein-1 Gene.** *Dis Colon Rectum* 2007.
11. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, Lothe RA: **ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis.** *Cell Oncol* 2006, **28:**259-272.
12. Lind GE, Ahlquist T, Lothe RA: **DNA hypermethylation of MAL: a promising diagnostic biomarker for colorectal tumors.** *Gastroenterology* 2007, **132:**1631-1632.
13. Puertollano R, Martin-Belmonte F, Millan J, de Marco MC, Albar JP, Kremer L, Alonso MA: **The MAL proteolipid is necessary for normal apical transport and accurate sorting of the influenza virus hemagglutinin in Madin-Darby canine kidney cells.** *J Cell Biol* 1999, **145:**141-151.
14. Tugores A, Rubio T, Rancano C, Alonso MA: **A tandem array of Sp-1 sites and a reverse initiator element are both required for synergistic transcriptional activation of the T-cell-specific MAL gene.** *DNA Cell Biol* 1997, **16:**245-255.
15. Alonso MA, Weissman SM: **cDNA cloning and sequence of MAL, a hydrophobic protein associated with human T-cell differentiation.** *Proc Natl Acad Sci U S A* 1987, **84:**1997-2001.
16. Marazuela M, Alonso MA: **Expression of MAL and MAL2, two elements of the protein machinery for raft-mediated transport, in normal and neoplastic human tissue.** *Histol Histopathol* 2004, **19:**925-933.
17. Mimori K, Shiraishi T, Mashino K, Sonoda H, Yamashita K, Yoshinaga K, Masuda T, Utsunomiya T, Alonso MA, Inoue H, Mori M: **MAL gene expression in esophageal cancer suppresses motility, invasion and tumorigenicity and enhances apoptosis through the Fas pathway.** *Oncogene* 2003, **22:**3463-3471.

18. Meling GI, Lothe RA, Børresen AL, Hauge S, Graue C, Clausen OP, Rognum TO: **Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis.** *Br J Cancer* 1991, **64:**475-480.
19. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH: **Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I.** *Scand J Gastroenterol* 1999, **34:**414-420.
20. Grunau C, Clark SJ, Rosenthal A: **Bisulfite genomic sequencing: systematic investigation of critical experimental parameters.** *Nucleic Acids Res* 2001, **29:**E65-E65.
21. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB: **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.** *Proc Natl Acad Sci U S A* 1996, **93:**9821-9826.
22. Li LC, Dahiya R: **MethPrimer: designing primers for methylation PCRs.** *Bioinformatics* 2002, **18:**1427-1431.
23. Clark SJ, Harrison J, Paul CL, Frommer M: **High sensitivity mapping of methylated cytosines.** *Nucleic Acids Res* 1994, **22:**2990-2997.
24. Melki JR, Vincent PC, Clark SJ: **Concurrent DNA hypermethylation of multiple genes in acute myeloid leukemia.** *Cancer Res* 1999, **59:**3730-3740.
25. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP: **Tissue microarrays for high-throughput molecular profiling of tumor specimens.** *Nat Med* 1998, **4:**844-847.
26. Diep CB, Thorstensen L, Meling GI, Skovlund E, Rognum TO, Lothe RA: **Genetic tumor markers with prognostic impact in Dukes' stages B and C colorectal cancer patients.** *J Clin Oncol* 2003, **21:**820-829.
27. Lothe RA, Peltomäki P, Meling GI, Aaltonen LA, Nyström-Lahti M, Pylkkänen L, Heimdal K, Andersen TI, Møller P, Rognum TO, Fosså SD, Haldorsen T, Langmark F, Brøgger A, de la Chapelle A, Børresen AL: **Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history.** *Cancer Res* 1993, **53:**5849-5852.
28. Thorstensen L, Lind GE, Løvig T, Diep CB, Meling GI, Rognum TO, Lothe RA: **Genetic and Epigenetic Changes of Components Affecting the WNT Pathway in Colorectal Carcinomas Stratified by Microsatellite Instability.** *Neoplasia* 2005, **7:**99-108.
29. Martin-Belmonte F, Kremer L, Albar JP, Marazuela M, Alonso MA: **Expression of the MAL gene in the thyroid: the MAL proteolipid, a component of glycolipid-enriched membranes, is apically distributed in thyroid follicles.** *Endocrinology* 1998, **139:**2077-2084.
30. Marazuela M, Acevedo A, Adrados M, Garcia-Lopez MA, Alonso MA: **Expression of MAL, an integral protein component of the machinery for raft-mediated pical transport, in human epithelia.** *J Histochem Cytochem* 2003, **51:**665-674.
31. Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM, Meltzer SJ: **A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer.** *Gastroenterology* 2006, **131:**797-808.
32. 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.
33. Deng G, Peng E, Gum J, Terdiman J, Sleisenger M, Kim YS: **Methylation of hMLH1 promoter correlates with the gene silencing with a region-specific manner in colorectal cancer.** *Br J Cancer* 2002, **86:**574-579.
34. Hatta M, Nagai H, Okino K, Onda M, Yoneyama K, Ohta Y, Nakayama H, Araki T, Emi M: **Down-regulation of members of glycolipid-enriched membrane raft gene family, MAL and BENE, in cervical squamous cell cancers.** *J Obstet Gynaecol Res* 2004, **30:**53-58.
35. Osborn NK, Ahlquist DA: **Stool screening for colorectal cancer: molecular approaches.** *Gastroenterology* 2005, **128:**192-206.
36. Itzkowitz SH, Jandorf L, Brand R, Rabeneck L, Schroy PC III, Sontag S, Johnson D, Skoletsky J, Durkee K, Markowitz S, Shuber A: **Improved Fecal DNA Test for Colorectal Cancer Screening.** *Clin Gastroenterol Hepatol* 2007, **5:**111-117.
37. Kopelovich L, Crowell JA, Fay JR: **The epigenome as a target for cancer chemoprevention.** *J Natl Cancer Inst* 2003, **95:**1747-1757.

Paper II

Terje Ahlquist, Guro E Lind, Vera L Costa, Gunn I Meling, Morten Vatn, Geir S Hoff, Torleiv O Rognum, Rolf I Skotheim, Espen Thiis-Evensen and Ragnhild A Lothe.

Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers.

Molecular Cancer, In Press

Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers

Terje Ahlquist[1,2], Guro E Lind[1,2], Vera L Costa[1,3], Gunn I Meling[4,5,*], Morten Vatn[5,6], Geir S Hoff[7], Torleiv O Rognum[4,5], Rolf I Skotheim[1,2], Espen Thiis-Evensen[6], and Ragnhild A Lothe*[1,2,9]

[1]Department of Cancer Prevention, Institute for Cancer Research, Norwegian Radium Hospital, Rikshospitalet University Hospital, Oslo, Norway. [2]Centre for Cancer Biomedicine, University of Oslo, Norway. [3]Department of Genetics, Portuguese Oncology Institute, Porto, Portugal. [4]Institute of Forensic Medicine, Rikshospitalet Medical Centre, University of Oslo, Oslo, Norway. [5]Faculty of Medicine, University of Oslo, Norway. [6]Medical Department, Rikshospitalet Medical Centre, Oslo, Norway. [7]Department of Medicine, Division of Gastroenterology, Telemark Hospital, Skien, Norway. [9]Department of Molecular Biosciences, University of Oslo, Oslo, Norway.

Email: Terje Ahlquist – terje.c.ahlquist@rr-research.no; Guro E Lind – guro.elisabeth.lind@rr-research.no; Vera L Costa - veralmcosta@gmail.com; Gunn I Meling - gi@meling.net; Morten Vatn - morten.vatn@rikshospitalet.no; Geir S Hoff - geir.hoff@sthf.no; Torleiv O Rognum - t.o.rognum@medisin.uio.no; Rolf I Skotheim - rolf.i.skotheim@rr-research.no; Espen Thiis-Evensen - espen.thiis-evensen@rikshospitalet.no; Ragnhild A Lothe - rlothe@rr-research.no

* Corresponding author

Abstract

Background: Multiple epigenetic and genetic changes have been reported in colorectal tumors, but few of these have clinical impact. This study aims to pinpoint epigenetic markers that can discriminate between non-malignant and malignant tissue from the large bowel, i.e. markers with diagnostic potential.
The methylation status of eleven genes (*ADAMTS1*, *CDKN2A*, *CRABP1*, *HOXA9*, *MAL*, *MGMT*, *MLH1*, *NR3C1*, *PTEN*, *RUNX3*, and *SCGB3A1*) was determined in 154 tissue samples including normal mucosa, adenomas, and carcinomas of the colorectum. The gene-specific and widespread methylation status among the carcinomas was related to patient gender and age, and microsatellite instability status. Possible CIMP tumors were identified by comparing the methylation profile with microsatellite instability (MSI), *BRAF*-, *KRAS*-, and *TP53* mutation status.
Results: The mean number of methylated genes per sample was 0.4 in normal colon mucosa from tumor-free individuals, 1.2 in mucosa from cancerous bowels, 2.2 in adenomas, and 3.9 in carcinomas. Widespread methylation was found in both adenomas and carcinomas. The promoters of *ADAMTS1*, *MAL*, and *MGMT* were frequently methylated in benign samples as well as in malignant tumors, independent of microsatellite instability. In contrast, normal mucosa samples taken from bowels without tumor were rarely methylated for the same genes. Hypermethylated *CRABP1, MLH1*, *NR3C1*, *RUNX3,* and *SCGB3A1* were shown to be identifiers of carcinomas with microsatellite instability. In agreement with the CIMP concept, MSI and mutated *BRAF* were associated with samples harboring hypermethylation of several target genes.
Conclusion: Methylated *ADAMTS1*, *MGMT,* and *MAL* are suitable as markers for early tumor detection.

Background

Most cases of colorectal cancer (CRC) originate from adenomas. The malignant potential of adenomas increases with size, grade of dysplasia, and degree of villous components,[1] along with the number and order of genetic and epigenetic aberrations.[2] The majority (~85%) of the sporadic carcinomas are characterized by chromosomal aberrations, referred to as a chromosomal unstable (CIN) phenotype, whereas the smaller group (~15%) typically show microsatellite instability (MSI) caused by defect DNA mismatch repair.[2] Most CIN tumors are microsatellite stable (MSS). A third molecular phenotype characteristic to a subgroup of CRC is the CpG island methylator phenotype (CIMP).[3] CIMP-positive tumors display methylation of multiple loci, are associated with proximal location in the colon, and are often microsatellite unstable. *BRAF* mutations are restricted to CIMP positive tumors, which may be sub-classified according to a certain combination of epigenetic and genetic changes.[4]

Here we have compared the time of occurrence and co-variation of multiple epigenetic markers in normal colon samples with those of adenomas and carcinomas in order to pinpoint early onset markers for neoplastic transformation.

Materials and methods

Tissue samples

Included in the present study are twenty-one normal colon mucosa samples from twenty deceased, cancer-free individuals, median age 52.5, range 33-86 (called N1 henceforth); 18 normal colon mucosa samples (N2) from 18 CRC patients, median age 70.5, range 24-89 (taken at distance (>10cm) from the carcinoma); 63 adenomas, median size 8mm, range 5-50mm, from 52 individuals, median age 67, range 62-72; and 52 carcinomas from 51 patients, median age 70, range 33-92. The colon, including the rectum, is divided into proximal and distal sections; the proximal, or right side, spans from coecum to two thirds of the way across transversum; the distal, or left side, comprises the last third of the transversum, sigmoideum, and the rectum. This division originates from the primitive digestive tract, where the right side corresponds from the midgut, while the left side corresponds to the hindgut. The number of proximal *versus* distal samples in the series is as follows: N1 (10 *vs.* 11); N2 (7 *vs.* 11); adenomas (18 *vs.* 45); and carcinomas (17 *vs.* 35). The carcinomas included here are from a series evaluated to contain on average 84% tumor cells.[5] Nine of the N2 samples correspond to nine primary tumors analyzed here. Most of the normal colon samples (26/39) consisted of mucosa only, whereas the remaining ones were taken from the bowel wall. The adenomas were obtained from individuals attending a Norwegian colonoscopy screening program.[6] The carcinomas and the N2 samples are from a prospective series collected from 7 hospitals in the Oslo region of Norway.[5] The N1 samples were autopsy material collected by one of the authors.

The MSI status was determined by use of two mononucleotide markers, BAT25 and BAT26, and a panel of dinucleotide markers. Details regarding the assessment of MSI status are given in Additional file 1.

All samples belong to approved research biobanks and are part of research projects approved according to national guidelines (Biobank; registered at the Norwegian Institute of Public Health. Projects: Regional Ethics Committee and National Data Inspectorate).

DNA methylation analyses

DNA from all samples was bisulfite modified and subjected to methylation specific polymerase chain reaction (MSP) for each gene.[7,8] Two of the authors independently scored all samples and the methylation status of all positive samples was confirmed by a second, independent round of MSP. If any discrepancies appeared, a third round of analysis was performed. In line with consensus scoring procedures, we considered carcinomas with band intensities as strong as the positive control (++) as methylated [see Additional file 2] for the gene promoter in question, while the benign lesions and normal mucosa were scored as positive also when weakly methylated, *i.e.* (+).

For detailed MSP protocol, primer sequences, and scoring criteria see Additional file 1. Representative MSP results can be seen in Figure 1.

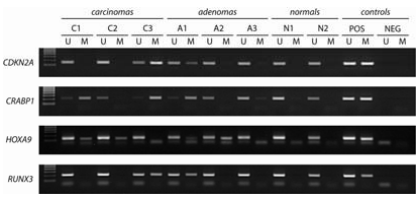


Figure 1. Representative methylation results in colorectal tumors and normal mucosa. Results of *CDKN2A*, *CRABP1*, *HOXA9*, and *RUNX3* in selected samples are shown. Positive controls (POS): NB, normal blood, for the unmethylated reaction and IVD, *in vitro* methylated DNA, for methylated reaction. Negative controls: dH_2O. U: unmethylated alleles, M: methylated alleles. The ladder (left lane) is the EZ Load™ 100bp Molecular Ruler (BioRad, Hercules, CA, USA).

Eleven genes, *ADAMTS1*, *CDKN2A* (encoding $p16^{INK4a}$), *CRABP1*, *HOXA9*, *MAL*, *MGMT*, *MLH1*, *NR3C1*, *PTEN*, *RUNX3,* and *SCGB3A1* (encoding HIN-1), were analyzed for promoter methylation by MSP. The methylation status of *ADAMTS1*, *CRABP1*, *MAL*, and *NR3C1* for the present series,[9,10] and the methylation status of *CDKN2A*, *MGMT*, and *MLH1* for the carcinomas [11] have previously been reported.

Quantitative MSP

Primers and probes for quantitative MSP (qMSP) were designed to specifically amplify fully methylated bisulfite-converted complementary sequences of the promoter of interest. The primers and probe sequences used for the *MGMT* [GenBank: NM_002412] are listed in Additional file 3. To normalize for DNA input in each sample, a reference gene (*ACTB* [12]) was used.

Fluorescence based real-time PCR assays were carried out in a reaction volume of 20 μL, consisting of 16.6mM ammonium sulphate; 67mM trizma preset; 6.7mM $MgCl_2$; 10mM mercaptoethanol; 0.1% DMSO; 200μM each of dATP, dCTP, dGTP, and dTTP; 600nM of each primer; 0.4 μL of Rox dye; 200nM of probe; 1 unit of platinum Taq polymerase (Invitrogen, Carlsbad, CA, USA), and 2 μl of bisulfite-modified DNA as a template. PCR was performed in separate wells for each primer/probe set and each sample was run in triplicate. Additionally, multiple water blanks were used, and as positive and negative control we used commercial methylated and unmethylated DNA (Millipore, Temecula, CA, USA). A series of dilutions of methylated DNA after bisulfite conversion were used for constructing a standard curve to quantify the amount of fully methylated alleles in each reaction. All amplifications were carried out in 96-well plates on an 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA), at 95ºC for 2 min followed by 45 cycles of 95ºC for 15 s, and 60ºC for 1 min.

In order to adjust for the possible various amounts of bisulfite treated DNA input in each PCR, the qMSP levels were normalized against the respective values of the internal reference gene (*ACTB*). The ratio thus generated constitutes an index of the percentage of input copies of DNA that are fully methylated at the primer- and probe-binding sites. The ratio was multiplied by 100 for easier tabulation (methylation level = target gene / reference gene x 100).

A given sample was considered positive for promoter hypermethylation when amplification was detected in at least 2 of the triplicates of the respective qMSP analysis. The qMSP threshold was determined by adjusting the best fit of the slope and R2, using the calibration curve.

Selection criteria for the 11 gene promoters analyzed in the present study

Some of the genes analyzed were known to be targeted through promoter methylation in cancer, including colorectal cancer (*SCGB3A1*, *RUNX3*, *CDKN2A*, *MLH1*, and *MGMT*). *HOXA9* was a potential new methylation target in colorectal cancer. *ADAMTS1*, *CRABP1*, *MAL*, and *NR3C1* were identified as novel epigenetically silenced target genes in colorectal cancer by our group [9,10,13] They were selected to be tested in combination with known methylated genes in a large series of colorectal lesions to check for interdependencies. The methylation status of all included genes was compared in a series of normal mucosa from individuals without cancer with those of normal, benign and malignant tissue from the large bowel of cancer patients. Only two previous studies have compared gene methylation among the same four types of sample groups as investigated here.[14,15] The first only investigated one gene and the latter 10 genes among which only three overlapping the present selected gene list.

Gene mutation status of BRAF, KRAS and TP53.

The present carcinoma series form a part of a series previously studied for genetic changes, including *BRAF, KRAS* and *TP53*.[16,17] The specific mutation status of the individual tumors included here can be found in Additional file 4.

Statistics

The 2x2 contingency tables were analyzed using Fisher's exact test and 3x2 tables were analyzed by the Pearson χ^2 test. Non-parametric analyses were performed using the Kruskal-Wallis and Mann-Whitney tests. An independent T-test was performed when comparing continuous normally distributed data with two groups. The bivariate correlation analysis was performed with Pearson's correlation. In order to determine age-specific methylation for the genes we used logistic regression analysis. All two-tailed *P*-values were derived from statistical tests using the SPSS15.0 software (SPSS, Chicago, IL, USA), and considered statistically significant at $P \leq 0.05$. The methylation heat-map was generated by average linkage hierarchical clustering and Pearson correlation distance measure, using the SpotFire DecisionSite®9.0 software.

Seven individuals had multiple polyps in the colon, and to exclude potential bias when analyzing patient data such as sex and age, one polyp from each individual was randomly selected for statistical analyses.

Results

MSI status of colorectal tumors

Two of sixty-three (3%) polyps displayed MSI. Both were large (≥10mm) and located in the proximal colon. The carcinomas were pre-selected according to MSI-status and 27/52 (52%) were MSI-positive.

DNA promoter methylation in normal mucosa, adenomas, and carcinomas

The results of the MSP analyses of all samples and each gene are summarized in Figure 2, Table 1, and Additional file 5. The mean number of genes methylated per sample was 0.4 for the N1 group, 1.2 for N2, 2.2 for adenomas, and 3.9 for carcinomas, and was significantly different among the groups using Kruskal-Wallis test; $P < 0.0001$ (mean rank N1, 10.2; N2, 17.3; adenomas, 23.1; and carcinomas, 31.4). Overall, 6/21 (29%) of the N1 samples, 9/18 (50%) of the N2 samples, 52/63 (83%) of the adenomas, and 48/52 (92%) of the carcinomas, were methylated in one or more of the eleven analyzed genes.

	ADAMTS1		*CDKN2A*		*CRABP1*		*HOXA9*		*MAL*		*MGMT*		*MLH1*		*NR3C1*		*RUNX3*		*SCGB3A1*	
	M	U	M	U	M	U	M	U	M	U	M	U	M	U	M	U	M	U	M	U
Sample-type																				
N1	0	21	0	21	0	21	4	17	1	20	2	19	0	21	0	21	0	21	1	20
N2	1	17	2	16	0	18	7	11	2	16	7	10	1	17	0	18	0	17	1	17
Adenoma	23	40	10	53	7	53	22	40	45	18	23	39	0	63	2	61	4	59	4	59
Carcinomas	36	15	17	35	25	25	12	38	41	9	21	31	11	41	13	37	16	31	9	40
MSI status																				
Carcinomas																				
MSI	19	8	10	17	22	5	7	20	21	6	11	16	11	16	12	15	16	10	8	18
MSS	17	7	7	18	3	20	5	18	20	2	10	15	0	25	1	22	0	21	1	22
P value	NS		NS		**<0.0001**		NS		NS		NS		**<0.0001**		**0.001**		**<0.0001**		**0.026**	

Table 1. Gene promoter methylation and microsatellite instability
Abbreviations: M, methylated samples; U, unmethylated samples; NS, not significant; N1: non-cancerous normal samples; N2: normal samples from cancer patients. MSI status data is listed for the individual polyp. In some cases a patient may have several polyps.

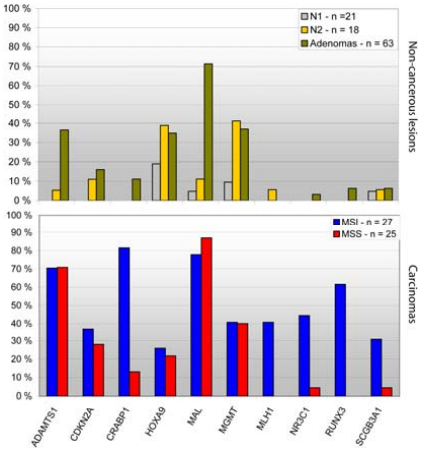


Figure 2. Methylation profiles of normal mucosa, adenomas, and carcinomas. Eleven genes were analyzed by MSP. Upper panel: non-cancerous lesions; lower panel: carcinomas stratified according to MSI-status. X-axis, the analyzed genes; Y-axis, the percentage of methylated samples. N1: normal colon samples from cancer-free individuals; N2: normal colon samples from cancer patients; MSI: microsatellite instability; MSS: microsatellite stability.

Statistically significant differences in methylation frequencies among sample groups were also evident at the single gene level. *ADAMTS1*, *CDKN2A*, *CRABP1*, *MLH1*, *NR3C1*, *RUNX3*, and *SCGB3A1* showed increasing methylation frequencies from adenomas to carcinomas, while *HOXA9*, *MAL*, and *MGMT* displayed overall equal methylation frequencies in all tumor subgroups. *PTEN* was unmethylated in carcinomas, and was thus not investigated in adenomas or included in the figures, tables (except Additional file 4) or statistics.

The more frequent promoter hypermethylation found among N2 samples compared with N1 samples was apparent both for the total number of methylated genes, and at the individual gene level (*MGMT*, P = 0.055). The reliability of our MSP scorings was tested by quantitative MSP analysis of one example gene performed in a blinded manner in another lab. The results were in perfect concordance [see Additional file 6].

No difference was seen in methylation frequencies between N2 samples with corresponding MSI-positive carcinomas (n = 6) and those with corresponding MSS carcinomas (n = 12).

Overall, gene methylation frequencies were higher among MSI than among MSS carcinomas, and were statistically significant for *CRABP1*, *MLH1*, *NR3C1*, *RUNX3*, and *SCGB3A1* ($P \leq 0.0001$, $P \leq 0.0001$, $P = 0.001$, $P \leq 0.0001$, and $P = 0.03$, respectively). Methylation of these genes showed a strong association to proximal carcinoma location, demonstrating the close connection between high methylation levels, proximal location and MSI ($P \leq 0.0001$, $P \leq 0.0001$, $P \leq 0.001$, $P = 0.001$, and $P = 0.04$, respectively). Association to site was also seen for *HOXA9* in N1 samples ($P = 0.04$). *HOXA9* was also more frequently methylated among non-cancerous normal mucosa (n = 20) from older patients compared to younger patients, indicating age-specific methylation ($P = 0.025$). However, this was not confirmed among the larger group of carcinomas (n = 52).

Interdependence among hypermethylated genes
From bivariate correlation analysis [see Additional file 7], methylation of *MLH1* was correlated with methylation of *CRABP1* (correlation coefficient 0.51; $P = 5x10^{-11}$), *NR3C1* (correlation coefficient 0.72; $P = 1x10^{-25}$) and *RUNX3* (correlation coefficient 0.57; $P = 6x10^{-14}$). Methylation of *RUNX3* itself was strongly correlated to methylation of both *NR3C1* (correlation coefficient 0.75; $P = 5x10^{-28}$) and *CRABP1* (correlation coefficient 0.67; $P = 3x10^{-20}$). Methylation of *NR3C1* and *CRABP1* was also correlated (correlation coefficient 0.59; $P = 4x10^{-15}$), as well as *ADAMTS1* and *MAL* (correlation coefficient 0.53; $P = 2x10^{-12}$).

Hierarchical clustering of samples according to gene methylation status showed that *MLH1* and *NR3C1* were most closely related, followed by *RUNX3* and *CRABP1*. In contrast, *HOXA9* and *MGMT* displayed methylation patterns independent from each other and the other genes (Figure 3).

Widespread methylation
Several samples harbored simultaneous promoter methylation of two or more of the analyzed genes [see Additional file 8]. The distribution of methylated gene numbers per sample did not appear to be bimodal. Neither N1 nor N2 samples displayed methylation of five or more genes, here denoted widespread methylation. Seven of 63 (11%) adenomas displayed widespread methylation, and these were by far larger in size (mean = 19 mm)

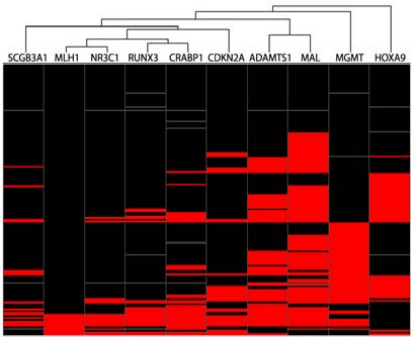


Figure 3. Methylation HeatMap. Hierarchical clustering reveals that methylation of *NR3C1* and *RUNX3* are most closely related, followed by *MLH1* and *CRABP1*. Methylation of *MGMT* and *HOXA9* are most independent both from each other and from rest of the set. The genes are presented in columns, while the samples are presented in rows. Black, unmethylated; red, methylated; and grey, missing values.

than the remaining adenomas (mean = 10 mm; P = 0.013). In carcinomas, widespread methylation was seen more frequently in MSI (16/27; 59%) than in MSS (3/25; 12%) samples (P = 0.001). All sixteen MSI samples with widespread methylation showed similar molecular profiles when DNA methylation status, *TP53*-, *KRAS*-, and *BRAF*-mutation status were considered, in line with a CIMP positive phenotype (Figure 4). The three MSS samples with widespread methylation included one tumor with *TP53* mutation, one with both *TP53* and *KRAS* mutation and one with *BRAF* mutation.

The distribution of the carcinomas combined with information regarding sex, age, MSI-status, and widespread methylation is illustrated in Figure 5. From the figure we see that widespread methylation is associated with proximal tumors derived from elderly women.

Discussion

We demonstrate in the present study aberrant promoter methylation of several genes, at variable frequencies, in the stepwise development of colorectal tumors.

An association between hypermethylation and lack of expression has previously been shown for all genes analyzed in the present study.[8-10,18-21] Although multiple genes are methylated in a cancer, only some are functionally involved in tumorigenesis,[22-24] whereas others with unknown functional contribution still may serve as good biomarkers from a diagnostic perspective.

Comparing methylation profiles of normal mucosa, adenomas and carcinomas of the large bowel

The identified methylation profiles of normal colorectal tissues, adenomas, and carcinomas demonstrated a stepwise increase in CpG island promoter methylation towards malignancy, indicating that their inactivation plays a role in the progression of the tumor. This was evident both for widespread methylation and at the single gene level (increasing frequencies of methylation from benign to malignant stages) with the exception of *HOXA9*, *MAL*, and *MGMT*. The lack of increase in methylation frequencies between non-malignant adenomas and carcinomas for these three genes may suggest that they are more important in the initiation of cancer, rather than in progression. These genes in addition to *ADAMTS1* were also hypermethylated in comparable frequencies among MSS and MSI carcinomas. These observations, and the fact that the separation of the MSI- and MSS-pathway is thought to occur early in colorectal tumorigenesis suggest that alterations of the four genes represent early events. *ADAMTS1* is believed to be an inhibitor of both angiogenesis and endothelial proliferation,[25] features commonly activated in cancer, as a tumor must turn on angiogenesis in order to grow larger than 1-2mm^3[26]. Members of the HOX gene family are shown to be commonly altered in several cancers, and to the best of our knowledge, this is the first report of *HOXA9* methylation in colorectal neoplasms. *HOXA9* methylation has received increasing interest in recent time as it is included in the HOXA-cluster which harbors methylation over a larger area than just a single promoter, indicating that methylation may mimic genetic micro-deletions and turn off a cluster of genes rather than just one at the time, i.e. yet another example of long range epigenetic silencing.[27-29]. *MAL* is involved in T-cell differentiation, especially in the late or intermediate stages.[30] It is also involved in

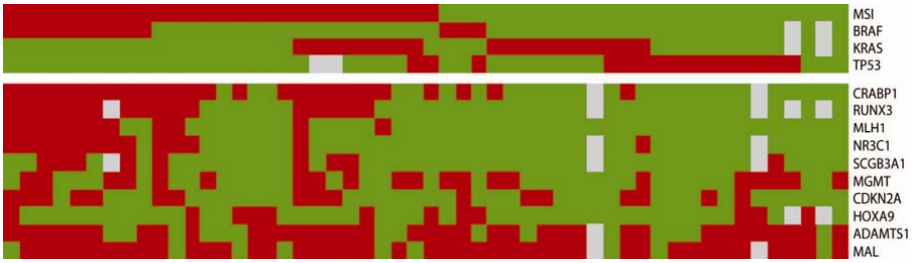


Figure 4. Genetic and epigenetic changes in colorectal carcinomas with known microsatellite status
The results are visualized according to genetic (top part of the figure) and epigenetic changes (lower part of the figure). The results are organized according to MSI, followed by *BRAF*-, *KRAS*-, *TP53*- and methylation-status of the MSI associated genes.

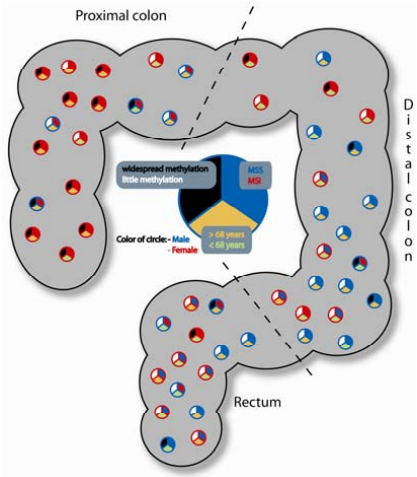


Figure 5. Distribution of colorectal carcinomas according to site associates with sex, age, MSI-status and methylation frequencies. The circles indicate 52 carcinomas placed according to site, the red circle=female, the blue=male. Top right section of the circle: blue=MSS, red=MSI. The lowest section: green=patient <68 years of age, yellow=patient ≥68 years. Widespread methylation is given in the top left section: white= methylation in < 5 genes, black=widespread methylation ≥ 5 genes.

polarization of epithelial cells caused by apical transport of lipids and proteins. Loss of cell polarity is often seen in neoplastic transformation.[31] For *MGMT* the early involvement is further supported by the fact that promoter methylation has previously been identified in aberrant crypt foci.[32]

Our data do not suggest that any of the markers included here were methylated in an age dependent manner. Of the 11 analyzed genes, six were unmethylated in all normal samples from non-affected individuals, excluding them as age-specific methylation targets. For two genes (*SCGB3A1* and *MAL*) only one of 21 samples was methylated. Although the sample in question was from an older individual (75 years), the resulting overall methylation frequency was only 5%. This is in strong contrast to the frequent reported age-specific methylation of the *N33* gene, which shows approximately 46% methylation among normal samples in general and 58% methylation in normal samples from individuals over 60 years.[33] *HOXA9* is the only gene in the present study harboring "frequent" promoter methylation in normal samples (19% overall, and 43% for individuals of 60 years or older). Binary regression analysis resulted in a significant *P* value, however, when using the same statistical analysis in the tumor sample series age dependence could not be confirmed. Both technical and biological aspects influence the interpretation of DNA promoter methylation analyses.

The importance of primer design is emphasized in the *PTEN* assay. Promoter hypermethylation of *PTEN* has been frequently reported in various tumor types, including CRC.[34-37] However, the majority of MSP primer sets used have failed to discriminate between *PTEN* and its frequently methylated pseudogene, leading to a high rate of false positives.[38] In the present study, we used MSP primers specifically designed to amplify the protein-encoding *PTEN* gene,[39] and showed that *PTEN* was not subject to promoter hypermethylation in colorectal carcinomas. A novel study confirms that methylation of *PTEN* is an unusual event in colorectal cancer as a whole.[40]

Interdependence among hypermethylated genes and wide spread methylation

The hierarchical clustering analysis of gene promoter methylation status in normal, benign, and malignant samples confirmed that the distribution of *HOXA9* and *MGMT* methylation frequencies across sample groups differed from the other genes. Overall, methylation of *NR3C1* and *RUNX3* had the highest correlation (figure 3 and Additional file 7), in addition to *MLH1,* which was also closely related to *NR3C1* and *RUNX3*. Furthermore, the present study confirmed that hypermethylation of *MLH1* was characteristic of right-sided sporadic colon tumors with MSI.[41] The lack of *MLH1* hypermethylation in adenomas analyzed in the present study supports the theory that CIMP and MSI-tumors arise from sessile serrated polyps rather than from adenomas.[42] *NR3C1*, *RUNX3*, *CRABP1*, and *SCGB3A1* were also shown to have the same characteristics as *MLH1*, supporting the hypothesis that DNA methylation plays a more prominent role in proximal than in distal carcinogenesis. *CRABP1*, *MLH1*, *NR3C1*, and *RUNX3* have recently been shown to belong to a panel of epigenetically regulated genes which best discriminate between CIMP-positive and CIMP-negative tumors, a phenotype strongly related with MSI status.[43]

We found that the MSI positive samples with V600E *BRAF* mutations were accompanied by promoter hypermethylation of several genes, in agreement with the CIMP phenotype (Figure 4). Furthermore, we also confirmed that MSS tumors with *TP53* mutations had less overall methylation, and thus in agreement with a CIMP negative phenotype. *KRAS* mutations were evenly distributed between MSI and MSS samples but seemingly the KRAS/MSI samples had more methylation than KRAS/MSS samples. Interestingly, three MSS samples had *BRAF* mutations, and all differed from the V600E mutation found among the MSI tumors.

Methylation markers suitable for early tumor detection

For genes previously analyzed for promoter methylation in normal colon samples, our results are within the expected range (*CDKN2A*, 0-33% (range of samples 9-100, total methylation frequency ~4%)[44-57]; *MGMT*, 0-39% (range of samples 12-220, total methylation frequency ~7%)[14,15,44,49,50,53,56-61]; and *MLH1*, 0-50% (range of samples 8-100, total methylation frequency ~5%)).[44,46,49,50,52,53,55-57,62-67] *SCGB3A1* and *RUNX3* have previously been analyzed in only one study, and both were unmethylated in 57 normal samples.[48] The study showing the highest methylation frequency of *CDKN2A* and *MLH1* were biased towards normal

samples taken distant from MSI- and CIMP-positive tumors,[46] thus a higher degree of methylation might be expected.

A suitable, highly specific, biomarker should be unmethylated in normal mucosa from healthy individuals and frequently methylated in carcinomas, and possibly also in benign lesions. To date, only few such markers have been identified,[10,68,69] and one of the most suitable ones, Vimentin, is non-expressed in a normal, healthy, colon.[69] The fact that an important biomarker is non-expressed in normal tissue supports the choice of a low threshold for methylation positive early lesions, applied in the present search for early onset biomarkers. Hypermethylation of genes such as *ADAMTS1* and *MAL* are also suitable biomarkers for early detection, as they are infrequently methylated in normal mucosa taken from individuals without cancer (0% and 5%, respectively), but highly methylated in malignant lesions (71% and 82%, respectively)[9,13]. In addition, both are frequently hypermethylated among the adenomas (37% and 71%, respectively) independent of size. Of course, sufficient sensitivity and specificity of these hypermethylation markers must be shown in feces or blood samples for the purpose of non-invasive testing. It should be note that this is an obstacle yet to be overcomed by suggested markers in existing non-invasive tests.

It has been speculated that methylation of specific genes, such as *MGMT*, may yield a so-called "field effect", providing favorable conditions for further alterations which eventually might lead to tumor formation.[58,70] The initial steps in tumorigenesis might be due to an epigenetic disruption of a progenitor/stem cell which may be followed by genetic mutations of gatekeeper genes, and the subsequent acquisition of other genetic and epigenetic alterations.[71] This model provides a possible explanation of why we see relatively high methylation frequencies for genes such as *MGMT,* and *HOXA9* in normal samples taken from cancer patients.

Summarized, this study has shown that gene-specific promoter hypermethylation is an early event in colorectal tumorigenesis, exemplified by hypermethylation of *MGMT* in adenomas and normal mucosa from cancer patients, and by the high frequency of *ADAMTS1* and *MAL* methylation in polyps irrespective of size. These markers are suitable as part of a panel aiming at detecting early colorectal lesions, and possibly a field effect in a "labile" colon. In general, we saw that aberrant CpG island hypermethylation increased with malignancy. Finally, methylation of *CRABP1*, *MLH1*, *NR3C1*, *RUNX3*, and *SCGB3A1* were identifiers of MSI carcinomas.

Competing Interests

The author(s) declare that they have no competing interests.

Authors' Contributions

All authors have read and approved the final version of the manuscript. TA was main responsible for the laboratory analyses, performed statistical analyses, made all figures and drafted the manuscript. GEL participated in the study design, in experimental analyses and in the preparation of the manuscript. VLC performed the quantitative methylation specific PCR analysis. GIM collected the cancer series and provided the clinicopathological information. MV participated in the screening study from which we received adenomas and patient information. GSH was responsible for the screening study from which we received adenomas and patient information. TOR collected and provided normal mucosa from non-cancerous individuals, the carcinoma series and participated in scientific discussions. RIS contributed to the statistical analyses and in scientific discussions. ETE participated in the screening study from which we received adenomas and patient information as well as in study design and scientific discussions. RAL conceived the study, participated in the evaluation of the results and in manuscript preparation.

Acknowledgements

This study was funded by grants from the Norwegian Research Council (163962/V50 and 161448/V40, RAL), the latter supporting TA as a PhD student. Funding was also received from the Norwegian Cancer Society (A95068, RAL) supporting GEL as post doctoral fellow.

References

1. Ponz dL, Di Gregorio C: **Pathology of colorectal cancer.** *Dig Liver Dis* 2001, **33:**372-388.

2. Grady WM, Markowitz SD: **Genetic and epigenetic alterations in colon cancer.** *Annu Rev Genomics Hum Genet* 2002, **3:101-28. Epub@2002 Apr 15.:**101-128.

3. 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.

4. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K et al.: **Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer.** *Proc Natl Acad Sci U S A* 2007, **104:**18654-18659.

5. Meling GI, Lothe RA, Børresen AL, Hauge S, Graue C, Clausen OP, Rognum TO: **Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis.** *Br J Cancer* 1991, **64:**475-480.

6. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH: **Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I.** *Scand J Gastroenterol* 1999, **34:**414-420.

7. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL: **A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA**

strands. *Proc Natl Acad Sci U S A* 1992, **89:**1827-1831.

8. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB: **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.** *Proc Natl Acad Sci U S A* 1996, **93:**9821-9826.

9. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, Lothe RA: **ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis.** *Cell Oncol* 2006, **28:**259-272.

10. Lind GE, Ahlquist T, Lothe RA: **DNA hypermethylation of MAL: a promising diagnostic biomarker for colorectal tumors.** *Gastroenterology* 2007, **132:**1631-1632.

11. Lind GE, Thorstensen L, Løvig T, Meling GI, Hamelin R, Rognum TO, Esteller M, Lothe RA: **A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines.** *Mol Cancer* 2004, **3:28**.

12. Eads CA, Lord RV, Wickramasinghe K, Long TI, Kurumboor SK, Bernstein L, Peters JH, DeMeester SR, DeMeester TR, Skinner KA et al.: **Epigenetic patterns in the progression of esophageal adenocarcinoma.** *Cancer Res* 2001, **61:**3410-3418.

13. Lind GE, Ahlquist T, Kolberg M, Berg M, Eknaes M, Alonso MA, Kallioniemi A, Meling GI, Skotheim RI, Rognum TO et al.: **Hypermethylated MAL gene - a silent marker of early colon tumorigenesis.** *J Transl Med* 2008, **6:13.:**13.

14. Nagasaka T, Goel A, Notohara K, Takahata T, Sasamoto H, Uchida T, Nishida N, Tanaka N, Boland CR, Matsubara N: **Methylation pattern of the O6-methylguanine-DNA methyltransferase gene in colon during progressive colorectal tumorigenesis.** *Int J Cancer* 2008, **122:**2429-2436.

15. Bai AH, Tong JH, To KF, Chan MW, Man EP, Lo KW, Lee JF, Sung JJ, Leung WK: **Promoter hypermethylation of tumor-related genes in the progression of colorectal neoplasia.** *Int J Cancer* 2004, **112:**846-853.

16. Ahlquist T, Bottillo I, Danielsen SA, Meling GI, Rognum TO, Lind GE, Dallapiccola B, Lothe RA: **RAS signaling in colorectal carcinomas through alteration of RAS, RAF, NF1, and/or RASSF1A.** *Neoplasia* 2008, **10:**680-6, 2.

17. Diep CB, Thorstensen L, Meling GI, Skovlund E, Rognum TO, Lothe RA: **Genetic tumor markers with prognostic impact in Dukes' stages B and C colorectal cancer patients.** *J Clin Oncol* 2003, **21:**820-829.

18. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG: **Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia.** *Cancer Res* 1999, **59:**793-797.

19. Krop IE, Sgroi D, Porter DA, Lunetta KL, LeVangie R, Seth P, Kaelin CM, Rhei E, Bosenberg M, Schnitt S et al.: **HIN-1, a putative cytokine highly expressed in normal but not cancerous mammary epithelial cells.** *Proc Natl Acad Sci U S A* 2001, **98:**9796-9801.

20. Alaminos M, Davalos V, Cheung NK, Gerald WL, Esteller M: **Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma.** *J Natl Cancer Inst* 2004, **96:**1208-1219.

21. Goel A, Arnold CN, Tassone P, Chang DK, Niedzwiecki D, Dowell JM, Wasserman L, Compton C, Mayer RJ, Bertagnolli MM et al.: **Epigenetic inactivation of RUNX3 in microsatellite unstable sporadic colon cancers.** *Int J Cancer* 2004, **112:**754-759.

22. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, Wright FA, Feramisco JD, Peltomaki P, Lang JC et al.: **Aberrant CpG-island methylation has non-random and tumour-type-specific patterns.** *Nat Genet* 2000, **24:**132-138.

23. Hayashi H, Nagae G, Tsutsumi S, Kaneshiro K, Kozaki T, Kaneda A, Sugisaki H, Aburatani H: **High-resolution mapping of DNA methylation in human genome using oligonucleotide tiling array.** *Hum Genet* 2007, **120:**701-711.

24. Schuebel KE, Chen W, Cope L, Glockner SC, Suzuki H, Yi JM, Chan TA, Van NL, Van CW, van den BS et al.: **Comparing the DNA hypermethylome with gene mutations in human colorectal cancer.** *PLoS Genet* 2007, **3:**1709-1723.

25. Kuno K, Bannai K, Hakozaki M, Matsushima K, Hirose K: **The carboxyl-terminal half region of ADAMTS-1 suppresses both tumorigenicity and experimental tumor metastatic potential.** *Biochem Biophys Res Commun* 2004, **319:**1327-1333.

26. Folkman J: **Angiogenesis.** *Annu Rev Med* 2006, **57:**1-18.

27. Novak P, Jensen T, Oshiro MM, Wozniak RJ, Nouzova M, Watts GS, Klimecki WT, Kim C, Futscher BW: **Epigenetic inactivation of the HOXA gene cluster in breast cancer.** *Cancer Res* 2006, **66:**10664-10670.

28. Strathdee G, Holyoake TL, Sim A, Parker A, Oscier DG, Melo JV, Meyer S, Eden T, Dickinson AM, Mountford JC et al.: **Inactivation of HOXA genes by hypermethylation in myeloid and lymphoid malignancy is frequent and associated with poor prognosis.** *Clin Cancer Res* 2007, **13:**5048-5055.

29. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ: **Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band.** *Nat Genet* 2006, **38:**540-549.

30. Alonso MA, Weissman SM: **cDNA cloning and sequence of MAL, a hydrophobic protein associated with human T-cell differentiation.** *Proc Natl Acad Sci U S A* 1987, **84:**1997-2001.

31. Marazuela M, Alonso MA: **Expression of MAL and MAL2, two elements of the protein machinery for raft-mediated transport, in normal and neoplastic human tissue.** *Histol Histopathol* 2004, **19:**925-933.

32. Chan AO, Broaddus RR, Houlihan PS, Issa JP, Hamilton SR, Rashid A: **CpG island methylation in aberrant crypt foci of the colorectum.** *Am J Pathol* 2002, **160:**1823-1830.

33. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP: **Aging and DNA methylation in colorectal mucosa and cancer.** *Cancer Res* 1998, **58:**5489-5494.

34. Kang YH, Lee HS, Kim WH: **Promoter methylation and silencing of PTEN in gastric carcinoma.** *Lab Invest* 2002, **82:**285-291.

35. Baeza N, Weller M, Yonekawa Y, Kleihues P, Ohgaki H: **PTEN methylation and expression in glioblastomas.** *Acta Neuropathol (Berl)* 2003, **106:**479-485.

36. Khan S, Kumagai T, Vora J, Bose N, Sehgal I, Koeffler PH, Bose S: **PTEN promoter is methylated in a proportion of invasive breast cancers.** *Int J Cancer* 2004, **112:**407-410.

37. Goel A, Arnold CN, Niedzwiecki D, Carethers JM, Dowell JM, Wasserman L, Compton C, Mayer RJ, Bertagnolli MM, Boland CR: **Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers.** *Cancer Res* 2004, **64:**3014-3021.

38. Zysman MA, Chapman WB, Bapat B: **Considerations when analyzing the methylation status of PTEN tumor suppressor gene.** *Am J Pathol* 2002, **160:**795-800.

39. Garcia JM, Silva J, Pena C, Garcia V, Rodriguez R, Cruz MA, Cantos B, Provencio M, Espana P, Bonilla F: **Promoter methylation of the PTEN gene is a common molecular change in breast cancer.** *Genes Chromosomes Cancer* 2004, **41:**117-124.

40. Goel A, Nagasaka T, Arnold CN, Inoue T, Hamilton C, Niedzwiecki D, Compton C, Mayer RJ, Goldberg R, Bertagnolli MM et al.: **The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer.** *Gastroenterology* 2007, **132:**127-138.

41. Jass JR: **Colorectal cancer: a multipathway disease.** *Crit Rev Oncog* 2006, **12:**273-287.

42. Jass JR: **Molecular heterogeneity of colorectal cancer: Implications for cancer control.** *Surg Oncol* 2007, **16 Suppl 1:S7-9. Epub;%2007 Nov 26.:**S7-S9.

43. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D et al.: **CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer.** *Nat Genet* 2006, **38:**787-793.

44. Lee S, Hwang KS, Lee HJ, Kim JS, Kang GH: **Aberrant CpG island hypermethylation of multiple genes in colorectal neoplasia.** *Lab Invest* 2004, **84:**884-893.

45. Ebert MP, Mooney SH, Tonnes-Priddy L, Lograsso J, Hoffmann J, Chen J, Rocken C, Schulz HU, Malfertheiner P, Lofton-Day C: **Hypermethylation of the TPEF/HPP1 gene in primary and metastatic colorectal cancers.** *Neoplasia* 2005, **7:**771-778.

46. Kawakami K, Ruszkiewicz A, Bennett G, Moore J, Grieu F, Watanabe G, Iacopetta B: **DNA hypermethylation in the normal colonic mucosa of patients with colorectal cancer.** *Br J Cancer* 2006, **94:**593-598.

47. Minoo P, Baker K, Goswami R, Chong G, Foulkes W, Ruszkiewicz A, Barker M, Buchanan D, Young J, Jass JR: **Extensive DNA methylation in normal colorectal mucosa in hyperplastic polyposis.** *Gut* 2006, **55:**1467-1474.

48. Takahashi T, Shigematsu H, Shivapurkar N, Reddy J, Zheng Y, Feng Z, Suzuki M, Nomura M, Augustus M, Yin J et al.: **Aberrant promoter methylation of multiple genes during multistep pathogenesis of colorectal cancers.** *Int J Cancer* 2006, **118:**924-931.

49. Derks S, Postma C, Moerkerk PT, van den Bosch SM, Carvalho B, Hermsen MA, Giaretti W, Herman JG, Weijenberg MP, de Bruine AP et al.: **Promoter methylation precedes chromosomal alterations in colorectal cancer development.** *Cell Oncol* 2006, **28:**247-257.

50. Ye C, Shrubsole MJ, Cai Q, Ness R, Grady WM, Smalley W, Cai H, Washington K, Zheng W: **Promoter methylation status of the MGMT, hMLH1, and CDKN2A/p16 genes in non-neoplastic mucosa of patients with and without colorectal adenomas.** *Oncol Rep* 2006, **16:**429-435.

51. Lee M, Sup HW, Kyoung KO, Hee SS, Sun CM, Lee SN, Koo H: **Prognostic value of p16(INK4a) and p14(ARF) gene hypermethylation in human colon cancer.** *Pathol Res Pract* 2006, **202:**415-424.

52. Park SJ, Rashid A, Lee JH, Kim SG, Hamilton SR, Wu TT: **Frequent CpG island methylation in serrated adenomas of the colorectum.** *Am J Pathol* 2003, **162:**815-822.

53. Kim HC, Lee HJ, Roh SA, Kim JS, Yu CS, Kim JC: **CpG island methylation in familial colorectal cancer patients not fulfilling the Amsterdam criteria.** *J Korean Med Sci* 2008, **23:**270-277.

54. Ishiguro A, Takahata T, Saito M, Yoshiya G, Tamura Y, Sasaki M, Munakata A: **Influence of methylated p15 and p16 genes on clinicopathological features in colorectal cancer.** *J Gastroenterol Hepatol* 2006, **21:**1334-1339.

55. Deng G, Kakar S, Tanaka H, Matsuzaki K, Miura S, Sleisenger MH, Kim YS: **Proximal and distal colorectal cancers show distinct gene-specific methylation profiles and clinical and molecular characteristics.** *Eur J Cancer* 2008, **44:**1290-1301.

56. Belshaw NJ, Elliott GO, Foxall RJ, Dainty JR, Pal N, Coupe A, Garg D, Bradburn DM, Mathers JC, Johnson IT: **Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa.** *Br J Cancer* 2008, **99:**136-142.

57. 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.

58. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG et al.: **MGMT promoter methylation and field defect in sporadic colorectal cancer.** *J Natl Cancer Inst* 2005, **97:**1330-1338.

59. Cheng YW, Shawber C, Notterman D, Paty P, Barany F: **Multiplexed profiling of candidate genes for CpG island methylation status using a flexible PCR/LDR/Universal Array assay.** *Genome Res* 2006, **16:**282-289.

60. Huang ZH, Li LH, Yang F, Wang JF: **Detection of aberrant methylation in fecal DNA as a molecular screening tool for colorectal cancer and precancerous lesions.** *World J Gastroenterol* 2007, **13:**950-954.

61. Menigatti M, Pedroni M, Verrone AM, Borghi F, Scarselli A, Benatti P, Losi L, Di GC, Schar P, Marra G et al.: **O6-methylguanine-DNA methyltransferase promoter hypermethylation in colorectal carcinogenesis.** *Oncol Rep* 2007, **17:**1421-1427.

62. Koinuma K, Kaneda R, Toyota M, Yamashita Y, Takada S, Choi YL, Wada T, Okada M, Konishi F, Nagai H et al.: **Screening for genomic fragments that are methylated specifically in colorectal carcinoma with a methylated MLH1 promoter.** *Carcinogenesis* 2005, **26:**2078-2085.

63. Nuovo GJ, Nakagawa H, Sotamaa K, Chapelle AdL: **Hypermethylation of the MLH1 promoter with**

concomitant absence of transcript and protein occurs in small patches of crypt cells in unaffected mucosa from sporadic colorectal carcinoma. *Diagn Mol Pathol* 2006, **15:**17-23.

64. Ide T, Kitajima Y, Ohtaka K, Mitsuno M, Nakafusa Y, Miyazaki K: **Expression of the hMLH1 gene is a possible predictor for the clinical response to 5-fluorouracil after a surgical resection in colorectal cancer.** *Oncol Rep* 2008, **19:**1571-1576.

65. Hitchins MP, Lin VA, Buckle A, Cheong K, Halani N, Ku S, Kwok CT, Packham D, Suter CM, Meagher A et al.: **Epigenetic inactivation of a cluster of genes flanking MLH1 in microsatellite-unstable colorectal cancer.** *Cancer Res* 2007, **67:**9107-9116.

66. 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.

67. Xiong Z, Wu AH, Bender CM, Tsao JL, Blake C, Shibata D, Jones PA, Yu MC, Ross RK, Laird PW: **Mismatch repair deficiency and CpG island hypermethylation in sporadic colon adenocarcinomas.** *Cancer Epidemiol Biomarkers Prev* 2001, **10:**799-803.

68. 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.

69. Chen WD, Han ZJ, Skoletsky J, Olson J, Sah J, Myeroff L, Platzer P, Lu S, Dawson D, Willis J et al.: **Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene.** *J Natl Cancer Inst* 2005, **97:**1124-1132.

70. Giovannucci E, Ogino S: **DNA methylation, field effects, and colorectal cancer.** *J Natl Cancer Inst* 2005, **97:**1317-1319.

71. Feinberg AP, Ohlsson R, Henikoff S: **The epigenetic progenitor origin of human cancer.** *Nat Rev Genet* 2006, **7:**21-33.

Additional material

Additional file 1:
File format: DOC
Title: Supplementary information
Description: Methodological details which are not crucial for the understanding of the work, as well as Additional figure legends.

Additional file 2:
File format: JPG
Title: Titration of methylated DNA template illustrates the scoring thresholds for the methylation-specific polymerase chain reaction.
Description: Determination of scoring thresholds is visualized by a titration series of the *RUNX3* gene.

Additional file 3:
File format: XLS
Title: PCR primers used for methylation-specific PCR and microsatellite instability analyses.
Description: Information on primers and PCR details

Additional file 4:
File format: XLS
Title: Genetic and epigenetic raw data.
Description: Raw data from all analyses are listed for each tumor included.

Additional file 5:
File format: XLS
Title: Summarized methylation raw data.
Description: Methylation frequencies are presented for each of the eleven analyzed genes in the 4 different sample types according to what methylation level they exhibited (strong, weak or no methylation).

Additional file 6:
File format: JPG
Title: Comparison between MSP and quantitative MSP in normal mucosa samples.
Description: Hypermethylation of was analyzed with both non-quantitative- and quantitative MSP. Here the results from each method are presented.

Additional file 7:
File format: XLS
Title: Correlation between methylated genes.
Description: A correlation table including all analyzed for promoter hypermethylation shown that genes commonly methylated in MSI tumors are highly correlated.

Additional file 8:
File format: JPG
Title: Widespread methylation among normal colorectal samples, adenomas, and carcinomas.
Description: A histogram showing the total number of methylated genes per sample in non-cancerous normal mucosa, normal mucosa taken in distance from a primary tumor, adenomas, and carcinomas stratified according to MSI status.

Paper III

Terje Ahlquist, Irene Bottillo, Stine A Danielsen, Gunn I Meling, Torleiv O Rognum, Guro E Lind, Bruno Dallapiccola and Ragnhild A Lothe.

RAS signaling in colorectal carcinomas through alterations of RAS, RAF, NF1 and/or RASSF1A.

Neoplasia. 2008. Jul;10(7):680-6

NEOPLASIA
www.neoplasia.com

Volume 10 Number 7 July 2008 pp. 680–686 680

RAS Signaling in Colorectal Carcinomas through Alterations of *RAS*, *RAF*, *NF1*, and/or *RASSF1A*[1]

Terje Ahlquist[*,†], Irene Bottillo[‡,§], Stine A. Danielsen[*,†], Gunn I. Meling[¶,#], Torleiv O. Rognum[#], Guro E. Lind[*,†], Bruno Dallapiccola[‡,§] and Ragnhild A. Lothe[*,†]

*Department of Cancer Prevention, Institute for Cancer Research, Norwegian Radium Hospital, Rikshospitalet University Hospital, Oslo, Norway; †Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway; ‡IRCCS-CSS, San Giovanni Rotondo and CSS-Mendel Institute, Rome, Italy; §Department of Experimental Medicine and Pathology, "Sapienza, University of Rome," Rome, Italy; ¶Surgical Department, Faculty Division Akershus University Hospital, University of Oslo, Oslo, Norway; #Institute of Forensic Medicine, Rikshospitalet-Radiumhospitalet Medical Centre, University of Oslo, Oslo, Norway

Abstract

More than half of all colorectal carcinomas are known to exhibit an activated mitogen-activated protein kinase pathway. The *NF1* gene, a negative regulator of KRAS, has not previously been examined in a series of colorectal cancer. In the present study, primary colorectal carcinomas stratified according to microsatellite instability status were analyzed. The whole coding region of *NF1* was analyzed for mutations using denaturing high-performance liquid chromatography and sequencing, and the copy number alterations of *NF1* were examined using multiple ligation-dependent probe amplification and real-time polymerase chain reaction. The mutational hot spots in *KRAS* and *BRAF* were sequenced, and promoter hypermethylation status of *RASSF1A* was assessed with a methylation-specific polymerase chain reaction. One sample had two missense mutations in *NF1*, whereas nine additional tumors had intronic mutations likely to affect exon splicing. Interestingly, 8 of these 10 tumors were microsatellite-unstable. Four other tumors showed a duplication of *NF1*. Mutations in *KRAS* and *BRAF* were mutually exclusive and were present at 40% and 22%, respectively. *RASSF1A* was hypermethylated in 31% of the samples. We show that the RAS signaling network is extensively dysregulated in colorectal carcinomas, because more than 70% of the tumors had an alteration in one or more of the four examined components.

Neoplasia (2008) 10, 680–686

Introduction

Colorectal cancer (CRC) is one of the leading causes of cancer-related deaths in the western world today, and at least 50% of CRCs are thought to have a dysregulation of the RAS-RAF-MEK-ERK pathway, also known as the mitogen-activated protein kinase (MAPK) pathway [1]. When activated, this pathway leads to increased proliferation and reduced apoptosis, two of six crucial abilities of a cancer cell, as described by Hanahan and Weinberg [2]. There are several components in this pathway, which, theoretically, could be affected in cancer, and some are known mutational targets in cancer such as *KRAS* and *BRAF*. *KRAS* has been widely established as an important

Abbreviations: CRC, colorectal cancer; MAPK, mitogen-activated protein kinase; MSI, microsatellite instability; MSS, microsatellite stable; dHPLC, denaturing high-performance liquid chromatography; MSP, methylation-specific PCR; MLPA, multiple ligation-dependent probe amplification

Address all correspondence to: Ragnhild A. Lothe, Department of Cancer Prevention, Institute for Cancer Research, Norwegian Radium Hospital, Montebello, 0310 Oslo, Norway. E-mail: ragnhild.a.lothe@rr-research.no

[1]This article refers to supplementary materials, which are designated by Tables W1 and W2 and are available online at www.neoplasia.com.

Received 19 February 2008; Revised 12 April 2008; Accepted 18 April 2008

DOI 10.1593/neo.08312

oncogene since its first mutational report in 1984 [3], and it is now known that it is mutated in 21% of all human sporadic cancers, including one third of CRCs [1].[2] *BRAF* was shown to be a mutational target in cancer 5 years ago [4], and 20% of all human cancers harbors a mutation, including an estimated 13% of colorectal carcinomas.[2] Another potential target of this pathway is the *NF1* gene, which encodes neurofibromatosis type 1, a GTPase-activating protein (GAP), governing hydrolysis of KRAS-GTP to KRAS-GDP [5], thereby functioning as a negative regulator of KRAS signaling. The *NF1* gene is approximately 280 kb in size and maps to chromosome 17q11.2. It contains 61 exons, with an 11- to 13-kb transcript and an open reading frame coding for 2818 amino acids. There are two catalytical domains in NF1, which are important for its function, namely, the cAMP/PKA domain comprising exons 11 to 17 and the RAS-GRD (RAS GAP-related domain) domain comprising exons 21 to 27a [6–8]. Neurofibromatosis type 1, a dominant disorder, is caused by mutations in *NF1*, but somatic mutations in this gene can also contribute to tumorigenesis. Since the first mutation report of the gene in 1992 showing that one colorectal tumor (of 22) was mutated in *NF1* [9], it has been speculated to play a role in colorectal tumorigenesis. However, due to the large size of the gene and the fact that there are no mutational hot spots, mutation analysis of *NF1* in tumors has been very scarce. *RASSF1* (Ras association domain family 1) gene maps at chromosome 3p21.3, and its isoform A (*RASSF1A*) has been found hypermethylated in 40% of lung tumors [10] and in a large variety of human cancers, including CRC [11,12]. As implied by its designation, RASSF1A is thought to interact with KRAS through a Ras association domain that alters its effects. RASSF1A has several effects, including promotion of apoptosis, cell cycle arrest, and maintenance of genomic stability, abilities typical of tumor suppressor genes. Some of these effects refer to the negative regulation of KRAS [13]. Its association to, and its effect on, KRAS is still not solved, although increasing evidence points to a direct binding between RASSF1A and farnesylated KRAS (reviewed in the study of Donninger et al. [11]). The *KRAS* and *BRAF* mutation status together with the alteration of other upstream components affecting the RAS signaling have been reported for other cancers [14], but only two previous studies have examined alterations in *KRAS*, *BRAF*, and *RASSF1A* in the same series of colorectal neoplasms [15,16], and independent of cancer type, no previous study has included a detailed analyses of the *NF1* gene.

To provide further insight into the role of MAP kinase signaling in CRC, we carried out the first comprehensive mutation analyses of the *NF1* gene in colorectal carcinomas in comparison with alterations of *BRAF*, *KRAS*, and *RASSF1A* in a sample series selected to include a comparable number of samples with and without the microsatellite instability phenotype.

Materials and Methods

Tissue Specimen

Sixty-five sporadic colorectal carcinomas from 64 patients with a mean age of 70 years (range 33–92 years), and an equal distribution of male–female were included in the present study. Twenty-nine samples displayed microsatellite instability (MSI), whereas 36 were microsatellite-stable (MSS). All tumors were nonfamilial as assessed by written questionnaires and cross check with the Norwegian Cancer Registry [17]. The colon, including the rectum, was divided into proximal and distal sections: the proximal, or right side, spans from cecum to two thirds of the way across transversum; the distal, or left side, comprises the last third of the transversum, sigmoideum, and the rectum. Of the 65 samples, 23 were located in the proximal colon and 42 were located in the distal colon. The carcinomas are from a prospective series collected from seven hospitals in the Southeast region of Norway during 1987–1989 and contain, on average, 84% tumor cells [18]. The tumors have been selected to achieve a consistently higher number of MSI-positive tumors compared to the normal distribution (15%). By stratifying the samples according to the MSI status, we ensured that any results associated with the MSI or MSS group would be detected.

NF1 *Mutation Screening — DNA Amplification and Denaturing High-Performance Liquid Chromatography Analysis*

Twenty-four representative CRC samples were analyzed for mutations in the *NF1* gene. These samples were selected to resemble the remaining series with regards to sex, age, tumor location, MSI status, and *KRAS* and *BRAF* mutation status. The 61 *NF1* gene exons were amplified in 61 polymerase chain reaction (PCR) fragments of 172 to 579 bp. The primers were generally positioned approximately 50 to 60 bp from the intron–exon boundary to allow the detection of splicing defects while minimizing intronic polymorphisms. In total, 19,843 bases were screened per sample to obtain the final mutation status. The dHPLC was carried out as previously published [19], with minor alterations in the PCR protocol and denaturing high-performance liquid chromatography (dHPLC) methods. For details concerning the dHPLC, please refer to Table W1.

In short, the initial PCR was carried out in 25 μl of reaction volumes and was cooled at room temperature for 60 minutes to yield heteroduplex formation. The identification of somatic *NF1* gene mutations was carried out with dHPLC on a 3500HT WAVE DNA fragment analysis system (Transgenomic, Crewe, UK) equipped with a DNASep column (Transgenomic). Polymerase chain reaction products were examined through a 5% linear acetonitrile gradient for heteroduplexes with a separation flow rate of 1.5 ml/min. Commercially available WAVE Optimized Buffers (A, B, and D; Transgenomic) and Syringe Solution (Transgenomic) were used to provide highly reproducible retention times with WAVE System instrumentation. Resolution temperatures and starting concentrations of buffer B for dHPLC analysis are reported in Table W1.

Sequencing

For each dHPLC abnormal elution profile, genomic DNA was reamplified with dHPLC primers and directly sequenced in both directions on a 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Forward and reverse sequences were analyzed and compared with the mRNA reference sequence and with the chromosome 17 genomic contig reference sequence (NM_000267). The first base (position +1) of the initiator methionine is taken as the start of the cDNA. All missense and splicing mutations detected were absent on 200 control chromosomes belonging to the unaffected subjects.

[2]Sanger Institute — Catalogue of Somatic Mutations in Cancer (COSMIC) Web site.

KRAS *and* BRAF *Mutation Screening*

The mutational hot spots of *KRAS* (exons 2 and 3) and *BRAF* (exons 11 and 15) were directly sequenced in both directions for all samples (n = 65) on an ABI PRISM 377 DNA Sequencer (Applied Biosystems) and an ABI PRISM 3730 DNA Sequencer (Applied Biosystems). All nucleotide numbers are based on the cDNA reference sequence (*BRAF*, GenBank Accession No. NM_004333; *KRAS*, GenBank Accession No. NM_004985). For primer details please see Table W1.

Methylation-Specific PCR of RASSF1A

Methylation-specific PCR (MSP) of *RASSF1A* were performed with published primers [20]. Polymerase chain reaction conditions were as follows: denaturation and enzyme activation at 95°C for 15 minutes; 35 cycles of 30 seconds of denaturation at 95°C, 30 seconds of annealing at 62°C, and 30 seconds of elongation at 72°C; final extension at 72°C for 7 minutes.

Human placental DNA (Sigma Chemical Co., St. Louis, MO) treated *in vitro* with *Sss*I methyltransferase (New England Biolabs Inc., Beverly, MA) was used as a positive control for MSP of methylated alleles, whereas DNA from normal lymphocytes was used as a control for unmethylated alleles. The PCR products were separated using a 2% agarose gel before individual visual scoring by two people. Methylated samples with intensity equal to, or higher than, the positive control were considered to be hypermethylated.

Multiple Ligation-Dependent Probe Amplification Analysis

Screening for *NF1* single- and multiexon deletions was carried out in 24 of the colorectal carcinomas using the SALSA P081/082 NF1 (version 04, 05-02-2005) multiple ligation-dependent probe amplification (MLPA) assay (MRC Holland, Amsterdam, The Netherlands), as instructed by the manufacturer and previously reported [21]. In brief, two probes in each exon were hybridized to the individual tumor DNA, followed by a ligation of the nick between the probes, and PCR amplification with 6-FAM–labeled universal primers. The amplified product was analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems), and the results were exported to Coffalyzer v.5 Software (MRC Holland). As controls, and in each experiment, we used five normal blood samples taken from healthy individuals who do not show NF1 phenotypic traits as determined by clinical evaluation. Furthermore, these controls are verified to have an unaltered *NF1*, both at the nucleotide and at the copy number level. A ratio of ~1 should be obtained if both alleles are present. A reduction or increase in the peak area values to <0.7 or >1.3 was considered an indication of a deletion or a gain, respectively. DNA samples showing a reduction or increase in the MLPA peak area according to the chosen threshold values were reanalyzed by MLPA, and only the samples showing consistent results between the two experiments were scored as deleted or gained.

Real-Time Polymerase Chain Reaction

The gene gains identified by MLPA were also confirmed with a TaqMan Real-Time PCR experiment using an ABI 7000 Sequence Detection System (Applied Biosystems). Two TaqMan probes mapping in *NF1* exons 25 and 28, respectively, were designed by File Builder 3.1 software (Applied Biosystems). These were amplified separately together with the endogenous control (RNaseP) in 96-well fast plates following the recommended protocol (Applied Biosystems). All samples were analyzed in parallel, and the mean value was used for data analysis. In cases where N-fold was below the maximum N-fold copy number observed among the nondeleted DNA used as negative controls, it was accepted that the test sample harbored one copy of the target gene. In cases where N-fold resulted above the minimum N-fold copy number observed among the nondeleted DNA, it was accepted that the test sample harbored two or more copies of the target gene.

Statistics

For this study, 2 × 2 contingency tables were analyzed using Fisher's exact test, whereas 3 × 2 tables were analyzed by the Pearson chi-square test. An independent t test was performed when comparing continuous normally distributed data between two groups. All P values were derived from statistical tests using the SPSS Version 15.0 software (SPSS, Chicago, IL), and considered statistically significant at $P \leq .05$.

Results

NF1

One of the 24 carcinomas contained two missense mutations (D1302Y and V2577G), the first located within the RAS-GRD domain.

In silico protein modeling showed that D1302Y has lost an exposed negative charge, which may be important in protein folding and in binding to other proteins. The V2577G most likely has no effect on the neurofibromin function. Additional nine tumors displayed intronic mutations in the range of 4 to 57 bases away from the intron–exon boundary (Table W2).

Using MLPA, we found that another 4 (17%) of 24 samples had a gain of parts or of the whole gene, also confirmed with real-time analysis (Table W2).

Comparison of the molecular results with clinicopathologic data showed that 8 of 10 samples with exonic or intronic alterations in *NF1* occurred in MSI-positive tumors (P = .047), whereas 3 of 4 duplications occurred in MSS tumors.

KRAS

Direct sequencing of exons 2 and 3 of *KRAS* revealed that 26 (40%) of 65 tumors harbored a mutation (Table W2). All but two mutations were missense mutations and occurred in codon 12, 13, or 61. These two were a 3-bp insertion (TGG) in exon 2 (c.49insTGG) and a 3-bp deletion in exon 3, codons 62 to 63 (c.184_189delGAG; Table W2). Furthermore, two of the tumors had two *KRAS* mutations each. One displayed both G12A and V14I mutations, and the second had both G12D and G13D.

Mutations in *KRAS* were seemingly more often present in MSS tumors than in MSI tumors, 69% *versus* 46% (P = .08).

BRAF

Mutational analysis of *BRAF* gene revealed that 14 (22%) of 64 samples harbored a mutation. Eleven of these were the typical V600E mutation; the remaining three were D594G, L597Q, and G1406C (Table W2).

Mutations in *BRAF* were strongly associated to MSI, female gender, and proximal location (P = .006, P = .015, and P = .025, respectively). Figure 1 illustrates the individual localization of each mutation in *KRAS*, *BRAF*, and *NF1*.

Figure 1. Site distribution of mutations within each gene. The mutations for the respective gene are placed according to their sequence position. In (a) and (b), only the exons in blue have been analyzed. In (c), all exons are analyzed, and the exons in orange indicate those that are only expressed in isoforms. To the right, representative sequencing results of mutant samples are presented.

RASSF1A

By MSP analysis, we found that 18 (31%) of 59 samples were hypermethylated in the promoter of *RASSF1A*. Methylation of the gene was more frequent in the distally located tumors ($P = .041$), but was not overlapping with the MSS phenotype. In eight tumors, hypermethylation of *RASSF1A* was the only observed alteration among the four genes analyzed here. We found no covariance between *RASSF1A* methylation and mutation status of either of the analyzed genes.

Dysregulation of RAS Signaling

When looking at concurrent mutations in individual tumors, we found that *KRAS* and *BRAF* were mutually exclusive because all *BRAF* mutations occurred in wild type *KRAS* tumors and *vice versa* ($P < .0001$). The sample with *NF1* missense mutations was MSI-positive, proximally located, and harbored a *BRAF* mutation. When including the intronic mutations in the number of *NF1* mutations, six of eight *BRAF* mutations occurred in *NF1*-mutated samples

(P = .03), overlapping with the MSI. Three of the four duplications found in the *NF1* locus occurred in tumors with wild type *BRAF* and *KRAS*. The remaining tumor had both a *KRAS* mutation and duplication.

The occurrence of *RASSF1A* hypermethylation in the presence of other mutations did not show any trends toward coexistence or mutually exclusive nature.

Taken together, we found that 74% (48/65) of the tumors had an overactive RAS signaling pathway due to change of at least one of the four analyzed components (one alteration in 37/48, two alterations in 10/48, and three alterations in 1/48). For the 24 samples submitted for complete analyses, the number of samples with at least one alteration was 19 (79%): one alteration was seen in 14 samples, two alterations were seen in 4 samples, and three alterations were seen in 1 sample. All samples and alterations are summarized in Figure 2 and Table W2.

Discussion

This is the first report with an extensive analysis of the role of *NF1* mutations in colorectal tumorigenesis. Previous mutation studies have only looked at a small number of samples, usually in a limited part of the gene, in the RAS–GAP domain. The initial mutational report on *NF1* showed that 1 of 22 colorectal adenocarcinomas harbored a mutation in the RAS–GAP domain using single-strand conformation polymorphism [9]. Another study of 10 colorectal cell lines and 4 sporadic tumors using protein truncation test disclosed mutations in the *NF1* coding region in four MSI cell lines (40%) and one MSI tumor (25%). Two of the cell lines had in fact two mutations each [22]. A recent study examined five hereditary nonpolyposis CRC patients for mutations in five exons and found a mutation in one (20%) of the patients who had a homozygous germline mutation of *MLH1* [23]. Moreover, loss of heterozygosity (LOH) at loci within the *NF1* gene have been shown in primary colorectal tumors (range, 14–57%) [24,25]. One of these studies also used real-time expression studies of *NF1* in 55 of the carcinomas and found an increased expression among tumors compared with normal colon tissue. In the COSMIC database [26], 79 carcinomas of the colorectum were apparently included among the *NF1* data, yielding a mutation frequency of 11%. However, seven of nine mutations reported were from one study including seven cell lines, leaving only two of the mutations occurring in sporadic primary tumors.

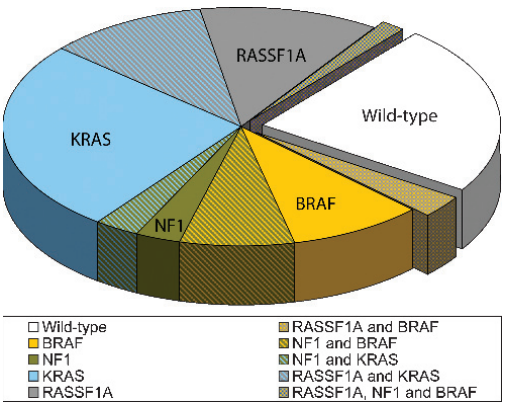


Figure 2. Alterations across the sample series. The pie chart indicates the four analyzed components and the percentage of tumors which showed alterations among these. Clockwise from the wild type pie, we see alterations in *RASSF1A* and *BRAF; BRAF; BRAF* and *NF1; NF1; NF1* and *KRAS; KRAS; KRAS* and *RASSF1A; RASSF1A; RASSF1A, NF1* and *BRAF*.

In this study, we found the *NF1* mutation profile to be in contrast to published germline mutation profile of NF1 patients[3] as well as to the somatic mutation profiles of malignant peripheral nerve sheath tumor taken from patients with and without the NF1 disease [27,28] (Bottillo I et al., unpublished observations). Furthermore, the median age of the patients included in the present CRC series is old, suggesting that potential NF1 carriers among them should have shown a debut of an NF1-associated cancer type. As no typical NF1-associated tumors are recorded, based on written questionnaires and confirmation of cancer diagnoses from the Norwegian Cancer Registry [17], it further support that the reported mutations are somatic. The observed intronic mutations prevailing among the colorectal tumors could be involved in alternative splicing but this remains to be elucidated. Four of the nine intronic mutations were indels of one or two bases in microsatellites and reflect replication slippage (which often occurs in such repetitive stretches of bases) left unrepaired by the defective DNA mismatch repair system [29]. No such indels were found in *KRAS* or *BRAF*.

Multiple ligation-dependent probe amplification results showed that 17% of the analyzed samples had gained parts of or whole of the *NF1* gene. This is not in accordance with the expectations of a tumor suppressor gene involved in tumorigenesis. A duplication of *NF1* could lead to a stronger negative regulation of KRAS, with subsequent stronger control of proliferation and differentiation. The duplications may arise as a consequence of the chromosomal instability present in three of the four tumors, which yield a wide range of gains and losses of whole or parts of chromosomes. As reported by Ĉaĉev et al. [30], colorectal tumors show a significant increase of *NF1* mRNA expression compared with corresponding normal tissue. They also showed that the expression of *NF1* isoform I (lacking exon 23a, located in the middle of the RAS-GRD domain) was significantly higher in tumor compared with normal tissue [30]. As of this, the present findings are in agreement with those of the study by Ĉaĉev et al. [30].

We also found a 40% mutation frequency of *KRAS*, which is within the expected range [26]. A mutated *KRAS* (in codons 12, 13, and 61) hinders the hydrolysis of GTP to GDP, and will keep KRAS in a constitutively active state, leading to phosphorylation of downstream effectors such as BRAF [31]. *BRAF* mutations are known to be strongly associated with MSI and CpG island methylator phenotype [32,33] and are found very often mutated in sessile-serrated adenomas, lesions often considered as a precursor of MSI-H tumors [34–38]. We found *BRAF* mutation in 22% of the samples, a higher frequency than in the mutation databases [26]. This reflects a bias due to the enrichment of MSI tumors in the present series. In one study, 71% of the MSI tumors had a V600E mutation in *BRAF*, as opposed to 7% in the chromosomal-unstable tumors [39], a figure comparable with the present series, as 18 (62%) of 29 of MSI tumors had *BRAF* mutations. The most common *BRAF* mutation, V600E, just as the common *KRAS* mutations, will lead to a constitutively active protein, as the activation loop of the protein is changed [31].

[3]NF1 International Mutation Database (http://www.nfmutation.org).

Some studies indicate an indirect interaction between RASSF1A and KRAS through RASSF5 (previously annotated NORE1A) [12], whereas others argue for a direct binding between RASSF1A and activated, farnesylated, KRAS [11]. Previous studies have also included *RASSF1A* when analyzing the impact of *KRAS* and *BRAF* mutations in colorectal tumorigenesis [13,15,16,40], and none of them found any co-occurrence between *RASSF1A* methylation and *BRAF* or *KRAS* mutation, in line with the present finding.

When adding the data of the fourth component, *NF1*, of the RAS signaling pathway, we found that more than 70% of the samples had a hyperactive RAS signaling. As the effect of RASSF1A on RAS signaling is still unclear, the eight samples with the sole alteration being hypermethylation may not be important for an overactive RAS signaling pathway. When we exclude the *RASSF1A* data from the combined analysis, 62% (40/65) of the samples had an overactive RAS signaling network, all due to *KRAS* or *BRAF* mutations, as the sample with the *NF1* missense mutations overlapped with *BRAF* mutation. If we include the *NF1* changes potentially affecting the splicing, 77% of the tumors have a dysregulation of the RAS signaling pathway.

In conclusion, we show that the RAS signaling network is extensively dysregulated in colorectal carcinomas as more than 70% of the tumors have an alteration in one or more of the four components.

References

[1] Fang JY and Richardson BC (2005). The MAPK signalling pathways and colorectal cancer. *Lancet Oncol* **6**, 322–327.

[2] Hanahan D and Weinberg RA (2000). The hallmarks of cancer. *Cell* **100**, 57–70.

[3] Santos E, Martin-Zanca D, Reddy EP, Pierotti MA, Della PG, and Barbacid M (1984). Malignant activation of a K-*ras* oncogene in lung carcinoma but not in normal tissue of the same patient. *Science* **223**, 661–664.

[4] Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, et al. (2002). Mutations of the *BRAF* gene in human cancer. *Nature* **417**, 949–954.

[5] Cichowski K and Jacks T (2001). *NF1* tumor suppressor gene function: narrowing the GAP. *Cell* **104**, 593–604.

[6] Martin GA, Viskochil D, Bollag G, McCabe PC, Crosier WJ, Haubruck H, Conroy L, Clark R, O'Connell P, and Cawthon RM (1990). The GAP-related domain of the *neurofibromatosis type 1* gene product interacts with *ras* p21. *Cell* **63**, 843–849.

[7] Xu GF, Lin B, Tanaka K, Dunn D, Wood D, Gesteland R, White R, Weiss R, and Tamanoi F (1990). The catalytic domain of the *neurofibromatosis type 1* gene product stimulates *ras* GTPase and complements *ira* mutants of *S. cerevisiae*. *Cell* **63**, 835–841.

[8] Fahsold R, Hoffmeyer S, Mischung C, Gille C, Ehlers C, Kucukceylan N, Abdel-Nour M, Gewies A, Peters H, Kaufmann D, et al. (2000). Minor lesion mutational spectrum of the entire *NF1* gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. *Am J Hum Genet* **66**, 790–818.

[9] Li Y, Bollag G, Clark R, Stevens J, Conroy L, Fults D, Ward K, Friedman E, Samowitz W, and Robertson M (1992). Somatic mutations in the *neurofibromatosis 1* gene in human tumors. *Cell* **69**, 275–281.

[10] Dammann R, Li C, Yoon JH, Chin PL, Bates S, and Pfeifer GP (2000). Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. *Nat Genet* **25**, 315–319.

[11] Donninger H, Vos MD, and Clark GJ (2007). The RASSF1A tumor suppressor. *J Cell Sci* **120**, 3163–3172.

[12] Hesson LB, Cooper WN, and Latif F (2007). The role of RASSF1A methylation in cancer. *Dis Markers* **23**, 73–87.

[13] Harada K, Hiraoka S, Kato J, Horii J, Fujita H, Sakaguchi K, and Shiratori Y (2007). Genetic and epigenetic alterations of Ras signalling pathway in colorectal neoplasia: analysis based on tumour clinicopathological features. *Br J Cancer* **97**, 1425–1431.

[14] McIntyre A, Summersgill B, Spendlove HE, Huddart R, Houlston R, and Shipley J (2005). Activating mutations and/or expression levels of tyrosine kinase receptors GRB7, RAS, and BRAF in testicular germ cell tumors. *Neoplasia* **7**, 1047–1052.

[15] Oliveira C, Velho S, Domingo E, Preto A, Hofstra RM, Hamelin R, Yamamoto H, Seruca R, and Schwartz S Jr (2005). Concomitant RASSF1A hypermethylation and KRAS/BRAF mutations occur preferentially in MSI sporadic colorectal cancer. *Oncogene* **24**, 7630–7634.

[16] Miranda E, Destro A, Malesci A, Balladore E, Bianchi P, Baryshnikova E, Franchi G, Morenghi E, Laghi L, Gennari L, et al. (2006). Genetic and epigenetic changes in primary metastatic and nonmetastatic colorectal cancer. *Br J Cancer* **95**, 1101–1107.

[17] Lothe RA, Peltomäki P, Meling GI, Aaltonen LA, Nystrom-Lahti M, Pylkkanen L, Heimdal K, Andersen TI, Moller P, and Rognum TO (1993). Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. *Cancer Res* **53**, 5849–5852.

[18] Meling GI, Lothe RA, Børresen AL, Hauge S, Graue C, Clausen OP, and Rognum TO (1991). Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis. *Br J Cancer* **64**, 475–480.

[19] De Luca A, Schirinzi A, Buccino A, Bottillo I, Sinibaldi L, Torrente I, Ciavarella A, Dottorini T, Porciello R, Giustini S, et al. (2004). Novel and recurrent mutations in the *NF1* gene in Italian patients with neurofibromatosis type 1. *Hum Mutat* **23**, 629.

[20] Koul S, Houldsworth J, Mansukhani MM, Donadio A, McKiernan JM, Reuter VE, Bosl GJ, Chaganti RS, and Murty VV (2002). Characteristic promoter hypermethylation signatures in male germ cell tumors. *Mol Cancer* **1**, 8.

[21] De Luca A, Bottillo I, Dasdia MC, Morella A, Lanari V, Bernardini L, Divona L, Giustini S, Sinibaldi L, Novelli A, et al. (2007). Deletions of *NF1* gene and exons detected by multiplex ligation-dependent probe amplification. *J Med Genet* **44**, 800–808.

[22] Wang Q, Montmain G, Ruano E, Upadhyaya M, Dudley S, Liskay RM, Thibodeau SN, and Puisieux A (2003). *Neurofibromatosis type 1* gene as a mutational target in a mismatch repair–deficient cell type. *Hum Genet* **112**, 117–123.

[23] Alotaibi H, Ricciardone MD, and Ozturk M (2008). Homozygosity at variant MLH1 can lead to secondary mutation in *NF1*, neurofibromatosis type I and early onset leukemia. *Mutat Res* **637**, 209–214.

[24] Cawkwell L, Lewis FA, and Quirke P (1994). Frequency of allele loss of DCC, p53, RBI, WT1, NF1, NM23 and APC/MCC in colorectal cancer assayed by fluorescent multiplex polymerase chain reaction. *Br J Cancer* **70**, 813–818.

[25] Leggett B, Young J, Buttenshaw R, Thomas L, Young B, Chenevix-Trench G, Searle J, and Ward M (1995). Colorectal carcinomas show frequent allelic loss on the long arm of chromosome 17 with evidence for a specific target region. *Br J Cancer* **71**, 1070–1073.

[26] Forbes S, Clements J, Dawson E, Bamford S, Webb T, Dogan A, Flanagan A, Teague J, Wooster R, Futreal PA, et al. (2006). COSMIC 2005. *Br J Cancer* **94**, 318–322.

[27] De Luca A, Buccino A, Gianni D, Mangino M, Giustini S, Richetta A, Divona L, Calvieri S, Mingarelli R, and Dallapiccola B (2003). *NF1* gene analysis based on DHPLC. *Hum Mutat* **21**, 171–172.

[28] Upadhyaya M, Han S, Consoli C, Majounie E, Horan M, Thomas NS, Potts C, Griffiths S, Ruggieri M, von Deimling A, et al. (2004). Characterization of the somatic mutational spectrum of the neurofibromatosis type 1 (*NF1*) gene in neurofibromatosis patients with benign and malignant tumors. *Hum Mutat* **23**, 134–146.

[29] Perucho M (1996). Cancer of the microsatellite mutator phenotype. *Biol Chem* **377**, 675–684.

[30] Ĉaĉev T, Radosevic S, Spaventi R, Pavelic K, and Kapitanovic S (2005). *NF1* gene loss of heterozygosity and expression analysis in sporadic colon cancer. *Gut* **54**, 1129–1135.

[31] Dhillon AS, Hagan S, Rath O, and Kolch W (2007). MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279–3290.

[32] Kawasaki T, Nosho K, Ohnishi M, Suemoto Y, Kirkner GJ, Fuchs CS, and Ogino S (2007). IGFBP3 promoter methylation in colorectal cancer: relationship with microsatellite instability, CpG island methylator phenotype, and p53. *Neoplasia* **9**, 1091–1098.

[33] Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, et al. (2007). Integrated genetic and epigenetic analysis

identifies three different subclasses of colon cancer. *Proc Natl Acad Sci USA* **104**, 18654–18659.

[34] Higuchi T and Jass JR (2004). My approach to serrated polyps of the colorectum. *J Clin Pathol* **57**, 682–686.

[35] Jass JR (2005). Serrated adenoma of the colorectum and the DNA-methylator phenotype. *Nat Clin Pract Oncol* **2**, 398–405.

[36] Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, Barker MA, Arnold S, McGivern A, Matsubara N, et al. (2004). *BRAF* mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut* **53**, 1137–1144.

[37] Makinen MJ (2007). Colorectal serrated adenocarcinoma. *Histopathology* **50**, 131–150.

[38] Wynter CV, Walsh MD, Higuchi T, Leggett BA, Young J, and Jass JR (2004). Methylation patterns define two types of hyperplastic polyp associated with colorectal cancer. *Gut* **53**, 573–580.

[39] Goel A, Nagasaka T, Arnold CN, Inoue T, Hamilton C, Niedzwiecki D, Compton C, Mayer RJ, Goldberg R, Bertagnolli MM, et al. (2007). The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer. *Gastroenterology* **132**, 127–138.

[40] Derks S, Postma C, Moerkerk PT, van den Bosch SM, Carvalho B, Hermsen MA, Giaretti W, Herman JG, Weijenberg MP, de Bruine AP, et al. (2006). Promoter methylation precedes chromosomal alterations in colorectal cancer development. *Cell Oncol* **28**, 247–257.

Table W1. Primer Sequences and dHPLC Conditions.

Exon	Primer FW	Primer RV	T_m PCR (°C)	Ampl. Size (bp)	tdHPLC	% B
1	5′-CTCCACAGACCCTCTCCTTG-3′	5′-GGACAGAGTAGGTGAGGGGA-3′	58	242	64–68	57–54
2	5′-AAACGTCATGATTTTCAATGGC-3′	5′-GGGGAATTTGCTTTCTTTTCTT-3′	58	281	55.5	56.6
3	5′-TTTCACTTTTCAGATGTGTGTTG-3′	5′-CTTTGTGAATTTGATCTTGAG-3′	58	210	55.5	54.8
4a	5′-GTTTGAAAATTTTCATAATAGAAA-3′	5′-CTCACAGCAGCTTTGACCTCC-3′	58	417	51–57	61–57
4b	5′-CAAGTGGTCCTCCTGCCTT-3′	5′-GTCAAAAACTAGTATCATGAATG-3′	58	283	55	55
4c	5′-TTTCCTAGCAGACAACTATCGA-3′	5′-ATTTGCTGTTGTTAGCATCCT-3′	62	308	54.5	57.5
5	5′-GAAGGAAGTTAGAAGTTTGTGACA-3′	5′-ATGGCTGGTAAGGATACGATTG-3′	62	172	54	51
6	5′-CATGTTTATCTTTTAAAAATGTTG-3′	5′-ATGTGAAGCAGTTTATTTTACTCAA-3′	62	332	54.5–56	58.6–56.6
7	5′-ATTTGCTATAATATTAGCTACATCTGG-3′	5′-GTTGATAAGTTCATAGGACTTGCTTT-3′	62	385	53–56.5	58–54.3
8	5′-GGATTTTACTGCCATTTGTGTG-3′	5′-TATCTAACTATATTTACTGATGCTGTTA-3′	58	276	56	56
9	5′-GCTGTTCTTTTTGGCTTC-3′	5′-CCAAAAGGTATTGCTAAATTAC-3′	58	183	54.5–56	52.2–49.7
9br	5′-GCTTAAAATTTGTATACAATAAAC-3′	5′-CCTGGAGTGGTGCTTCATGCAT-3′	58	193	55–60	52–48
10a	5′-CTACAGTGATAAACAGAGCAT-3′	5′-ATTCCTGCTGCTTTGGTT-3′	62	292	55–58	58–55
10a2	5′-CATTTTTTTGGTGTTTATGTATAGCAAG-3′	5′-GTGTATAGTTACCATTATAGTCACATC-3′	62	252	54–57–60.5	57–54–50
10b	5′-ATTATCCTGAGTCTTATGTC-3′	5′-TCTCAAAATTATCACACTAAGTTA-3′	58	229	54–57	56.2–51.2
10c	5′-ACCCTTTAGCAGTCACTGTC-3′	5′-CTGTGAGTAACAGGTAGATG-3′	58	307	54–59	59–54.4
11	5′-GAAAGAGCTCAATTTCTTAGC-3′	5′-CACTTCCAAAGGTTTTATGGT-3′	58	307	52–55	58–55
12a	5′-TGTATTCATTATGGGAGAATGCC-3′	5′-TGGAAGAATATTTGGAATGGTAAT-3′	58	269	54–56	56–54
12b	5′-GAGGTTTTTTAGGAGAGTCTC-3′	5′-ATGTGCTCTGTTTGTTTTCTG-3′	58	315	54–57.5	58–53.4
13	5′-CACAGTTTATTGCATTGTTAG-3′	5′-CTGCCTCAAAGCACATGGC-3′	62	380	57–61	59–54
14	5′-GCTCTTCCTACTCCTTTTGG-3′	5′-TATGCCCTTAGCAACAGAAA-3′	58	191	60	54
15	5′-ACTTGGCTGTAGCTGATTGA-3′	5′-TCAAGAGTCGCTCAGTAAAGT-3′	62	247	57	57
16	5′-CATTTTTTGTACTTTTGTCATGG-3′	5′-CTCTTATTTTTCACCTTTCTC-3′	58	579	55–58	63–60
17	5′-ATTTGGCTCTATGCCTGTGG-3′	5′-ACTGCACACAAACTAGGGTG-3′	58	385	55.5	58.8
18	5′-AGAAGTTGTGTACGTTCTTTTCT-3′	5′-GCGGTTATTGGTAGAAAGGAG-3′	58	367	53–56	57–54.4
19a	5′-TCATGTCACTTAGGTTATCTGG-3′	5′-CCTTCAAGTATTAGTGGGTTTTA-3′	58	242	55–57.5	58–55.7
19b	5′-TGAGGGGAAGTGAAAGAACT-3′	5′-GCAAAAAGCAAATAAAGCC-3′	58	236	53.5–57.5	56.5–52.5
20	5′-CCACCCTGGCTGATTATCG-3′	5′-GCATGTAAGAGAAGCAAAAATTA-3′	62	402	57–59	59–57
21	5′-AGCAAAAATTACTTCAGCAA-3′	5′-TCAGAGCCAGAAGAAAGATG-3′	58	393	57–59	59–57
22	5′-TGCTACTCTTTAGCTTCCTAC-3′	5′-GGCTGATTGTCTTCTTTTAAGG-3′	58	331	56.5–58	58.6–57
23.1	5′-TTTGTATCATTCATTTTGTGTGTA-3′	5′-CTTTTCACATAGAACCGCTGTTTTTT-3′	58	283	56–57	58.2–57.2
23.2	5′-GGCTTAATGTCTGTATA-3′	5′-GAGATTACCATTATTAATCTAAAGT-3′	58	270	53–59	57–51.3
23a	5′-AGCCAGAAATAGTAGACATGATTGGG-3′	5′-TCTACTAATTCTGGCACAAAATAG-3′	62	446	54.5	60.3
24	5′-TTGAACTCTTTGTTTTCATGTCTT-3′	5′-GATAATCTAGCTATCTTAAATTCC-3′	58	266	53–58	57–52.1
25	5′-AATTTATAGAATGAGGAATG-3′	5′-GTACCTGTTTTACATGAAGTTCCT-3′	54	335	52–54–57	58–56–53.7
26	5′-GCTTTGTCTAATGTCAAGTCA-3′	5′-GATAGTGAACACTCTCCGTTTAA-3′	62	342	56–58	58–54
27a	5′-ATGGTCCTGAGGTCTTTTTG-3′	5′-GCCACCAGGCCACTTGTTAG-3′	62	361	57	59
27b	5′-TTGCTTTTAAAATATTTTTTCATTTTAG-3′	5′-CCCAGTTGACTTAACAGGAATT-3′	58	330	55	55
28	5′-AAAATAAAATTGATTAGTGGCATCTG-3′	5′-AAATGTCACGTAAGGCTGTCG-3′	62	636	55–58	62–60
29	5′-TCTGGAGCCTTTTAGAATTTTATGT-3′	5′-TCAGTTTGATTTGGGGTTTGTTGC-3′	62	460	58–60.5	60–55.5
30	5′-GAAAAAATTTTGGAACTATAAGG-3′	5′-TAACAATTATTCTAAGAGAATTCAAAG-3′	58	322	51–56.5	58–54
31	5′-TTTTTTCCCCGAATTCTTTATG-3′	5′-CTTCAGAAAGCATGTAGACACTCAC-3′	58	425	55–57	61–59
32	5′-ATCTAGTATTTTTGAGGCCTCAG-3′	5′-CCTTCTGTACTATAGCATATCTG-3′	58	312	53–56	58–55
33	5′-TGCTAAAACTTTGAGTCCCATG-3′	5′-GTGCTCTAACACCAAGTTGC-3′	64	448	56–59	59–53.8
34	5′-TTCTAAATTCAAAATGAAACATGG-3′	5′-AAAAACACTTGCATGGACTG-3′	58	432	51.5–57	60–55
35	5′-GCATGGACTGTGTTATTGGTA-3′	5′-TCTGTGGATCTTTTAATTGCA-3′	58	319	53.5	56.8
36	5′-GCTGGACCAGTGGACAGAAC-3′	5′-GACGTTTAAATTTGAGGTCAATGA-3′	62	389	53–58	57.8–54.3
37	5′-TCCTGAATTCATTCCGAGATT-3′	5′-TCATTTTGGGTATCAGTGTTGAA-3′	58	237	54–56	55.5–53.5
38	5′-AACTGCAGTGTGTTTTGAAAGAG-3′	5′-GAGGTTCCTAGATTACTCAAATTTAG-3′	62	257	57–60	56–53
39	5′-TTGAACACAAAATTAAGTGAGCC-3′	5′-GAAGTAAGTTAGCCCTTATGTCTTAC-3′	62	318	56	55
40	5′-ATTCACATTCACATATGCATGTTTTACCTTC-3′	5′-CTTTGGTTCAAGACACTACAG-3′	62	547	55–56	61.6–61.1
41	5′-GTGCACATTTAACAGGTACTAT-3′	5′-ATCTAGAGATGGCCTAGGAAG-3′	62	373	55.5	58
42	5′-CTTGGAAGGAGCAAACGATGGTTG-3′	5′-CCATGTCAGTGTAGCAAAGTTTTTG-3′	58	356	55–60	56–52.2
43	5′-AGTGTATTCCCATTTATAGACACTG-3′	5′-CATTGAAAATAAGGTGGGAGA-3′	58	234	55–57.5	56–52.4
44	5′-GAAGTAACATTGAAATAGTTAGG-3′	5′-TCCAGTCTACTTTTAGGAGGCC-3′	58	271	58.5	55
45	5′-CATGAATAGGATACAGTCTTCTAC-3′	5′-GTTAAATGCTTACCCAGTAATGTG-3′	62	269	57	56
46	5′-CTCATCTCCCTTTAATTTTGGC-3′	5′-TCTGGAGAAGGATGGTTGATG-3′	58	295	54–56.5	57–55.1
47	5′-CTGTTACAATTAAAAGATACCTTG-3′	5′-GTATGCCTGCTTTAAGAACACACA-3′	62	185	55.5	51.4
48	5′-AAGGAAGAAAAATAGTAAATTAAGTCC-3′	5′-GTTTATAGCAAATTTTGCTCCTT-3′	58	423	53–58	61–56.9
48a	5′-ATTCAATAATTAAAACCAGATTCC-3′	5′-CTTTAGGAACTTGTAAAGCCACC-3′	58	327	54	58
49	5′-AGAATGTGTCCCCGTTGTTAA-3′	5′-TAATGAACCCATCCGGTTTG-3′	58	369	58.5	58.4
KRAS ex2	5′-ACTGGTGGAGTATTTGATA-3′	5′-GTATCAAAGAATGGTCCT-3′	50	—	—	—
KRAS ex3	5′-ATAATAGCCAATCCTAA-3′	5′-ATGGCATTAGCAAAG-3′	53	—	—	—
BRAF ex11	5′-TCATAATGCTTGCTCTGATAGGA-3′	5′-GGCCAAAAATTTAATCAGTGGA-3′	60	—	—	—
BRAF ex15	5′-TCCCTCTCAGGCATAAGGTAA-3′	5′-CGAACAGTGAATATTTCCTTTGAT-3′	58	—	—	—

T_m PCR, indicates PCR melting temperature (°C); *Ampl. Size*, amplicon size; *tdHPLC*, range in temperature used with high-performance liquid chromatography; *% B*, starting concentration for buffer B used in dHPLC.

Table W2. Detailed Somatic Events of Four Components in the MAPK Pathway.

Tumor ID	MSI Status	*KRAS*^{mut}	*BRAF*^{mut}	*NF1*^{mut}	MLPA	Real-time	*RASSF1A*
848	MSI	WT	WT	NP	NP		U
854	MSI	c.184-189delGAG	WT	NP	NP		U
884	MSI	WT	V600E	D1302Y/V2577G	WT		U
894	MSI	c.49insTTG	WT	NP	NP		U
910	MSI	WT	WT	c.(3114-50)delTG	WT		U
912	MSI	G13D	WT	NP	NP		U
955	MSI	WT	V600E	NP	NP		U
965	MSI	WT	V600E	NP	NP		U
980	MSI	WT	V600E	WT	WT		U
984	MSI	WT	V600E	WT	WT		U
988	MSI	WT	WT	NP	NP		ND
1022	MSI	WT	WT	NP	NP		U
1044	MSI	WT	V600E	c.480-57C>T	WT		U
1047	MSI	G12A/V14I	WT	WT	WT		M
1066	MSI	WT	WT	WT	WT		M
1117	MSI	WT	WT	NP	NP		U
1132	MSI	G12V	WT	NP	NP		U
1141	MSI	WT	WT	NP	NP		U
1190	MSI	WT	V600E	NP	NP		M
1193	MSI	WT	V600E	c.7395-7C>T	WT		U
1268	MSI	WT	V600E	Ex3+24G>A	WT		U
1273	MSI	WT	V600E	c.(1392+46_+53)delTT	WT		U
1314	MSI	WT	WT	NP	NP		U
1326	MSI	G13D	WT	c.(1392+46_+53)delT	WT		U
1341	MSI	WT	V600E	c.(61-4_-12)delT	WT		M
1349	MSI	WT	WT	NP	NP		M
1363	MSI	G13D	WT	WT	Gain of IVS27b-Ex49	1.66	U
1388A	MSI	WT	WT	NP	NP		M
1388C	MSI	G13D	WT	NP	NP		M
868	MSS	WT	WT	NP	NP		U
886	MSS	G12D	WT	NP	NP		U
887	MSS	G12C	WT	NP	NP		M
896	MSS	WT	WT	WT	Gain of whole gene	1.59	M
904	MSS	WT	WT	NP	NP		U
922	MSS	G12V	WT	NP	NP		U
923	MSS	G13D	WT	NP	NP		U
927	MSS	G12V	WT	NP	NP		M
946	MSS	WT	WT	NP	NP		U
948	MSS	G12R	WT	NP	NP		U
953	MSS	WT	WT	WT	WT		U
966	MSS	Q61L	WT	NP	NP		U
974	MSS	G12A	WT	NP	NP		U
976	MSS	G12D	WT	NP	NP		M
1013	MSS	WT	D594G	NP	NP		U
1024	MSS	G12C	WT	NP	NP		U
1027	MSS	G13D	WT	NP	NP		U
1029	MSS	G12D	WT	NP	NP		U
1046	MSS	WT	WT	NP	NP		U
1060	MSS	WT	WT	NP	NP		M
1069	MSS	WT	WT	NP	NP		U
1103	MSS	WT	WT	WT	Gain of whole gene	1.25	U
1111	MSS	WT	WT	NP	NP		M
1121	MSS	G12A	WT	WT	WT		M
1124	MSS	G12D/G13D	WT	WT	WT		ND
1166	MSS	G13D	WT	WT	WT		M
1167	MSS	WT	WT	WT	Gain of whole gene	1.67	M
1194	MSS	WT	WT	NP	NP		M
1197	MSS	WT	WT	NP	NP		ND
1287	MSS	WT	WT	c.2252-31A>G	WT		ND
1294	MSS	WT	G469R	NP	NP		U
1296	MSS	G13D	WT	WT	WT		U
1340	MSS	WT	WT	WT	WT		ND
1364	MSS	WT	L597Q	NP	NP		M
1369	MSS	G12D	ND	NP	NP		ND
1391	MSS	G12D	WT	c.2252-31A>T	WT		U

Paper IV

Terje Ahlquist, Ellen C Røyrvik, Marianne A Merok, Gunn I Meling, Annika Lindblom, Xiao-Feng Sun, Georgia Bardi, Matthias Kolberg, Arild Nesbakken and Ragnhild A Lothe.

Identification of RCC2 as a prognostic marker among multiple gene mutations in colorectal cancer with defect mismatch repair.

Manuscript

Identification of *RCC2* as a prognostic marker among multiple gene mutations in colorectal cancer with defect mismatch repair

Terje Ahlquist[1,2], Ellen C Røyrvik[1,2], Marianne A Merok[1,2,3], Gunn I Meling[4,5], Annika Lindblom[6], Xiao-Feng Sun[7], Georgia Bardi[8], Matthias Kolberg[1,2], Arild Nesbakken[3] and Ragnhild A Lothe[1,2,9]

[1]Department of Cancer Prevention, Institute for Cancer Research, Rikshospitalet-Radiumhospitalet Health Enterprise, Oslo, Norway. [2]Centre for Cancer Biomedicine, Faculty Division Norwegian Radium Hospital, University of Oslo, Oslo, Norway. [3]Department of Gastrointestinal Surgery, Aker University Hospital, University of Oslo, Oslo, Norway [4]Faculty of Medicine, University of Oslo, Norway. [5]Surgical Department, Faculty Division Akershus University Hospital, University of Oslo, Oslo, Norway. [6]Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden. [7]Department of Oncology, Institute of Clinical and Experimental Medicine, University of Linköping, Linköping, Sweden. [8]BioAnalytica-GenoType SA, Molecular Cytogenetic Research and Applications, Athens, Greece. [9]Department of Molecular Biosciences, University of Oslo, Oslo, Norway

ABSTRACT

Colorectal cancer (CRC) is one of the major cancer killers in the world, and any factor that better pinpoint patients with a poor prognosis may result in a prolonged life expectancy for a large number of cancer patients.

In order to identify novel prognostic markers a literature survey followed by a critical selection of tumorigenic genes were performed. We analyzed frameshift mutations in 41 genes with mostly coding mononucleotide repeats in two tumor series with a total of 202 microsatellite instable (MSI) colorectal tumors using capillary electrophoresis. The results were confirmed by sequencing in a random subset of the genes and tumors.

In total, the two series of MSI-tumors carried a mean number of 17 and 19 mutations, respectively. A strong association was seen between low mutation load and rectal location, both for individual genes (*ACVR2A*, *ASTE1*, *CASP5*, *MARCKS*, *MBD4*, *MRE11A*, *MSH3*, *TAF1B* and *TFGBR2)* as well as across all genes ($P = 0.008$).

Univariate survival analysis revealed that mutation in *RCC2* was associated with an increased five-year disease-free survival in both tumor series ($P = 0.035$ and 0.011, respectively). This finding was confirmed using multivariate analyses even with the inclusion of the strongest known predictor of prognosis to date, tumor stage at time of diagnosis ($P = 0.028$ and 0.021, respectively).

In conclusion, analysis of an (A)10 repeat in *RCC2* using readily available technology refines determination of prognosis in a group of MSI tumors. This may be useful in a future choice of treatment for stage II MSI-tumors.

INTRODUCTION

Microsatellite instability (MSI) is a phenotype seen in approximately 15% of all colorectal cancers (CRC).[1] It is one of three commonly described phenotypes in CRC alongside chromosomal instability (CIN), defined by large chromosomal rearrangements and aneuploidy, and CpG Island Methylator Phenotype (CIMP), which denotes tumors with promoter hypermethylation of a large number of genes.[1] The CIMP tumors show a large degree of overlap with the MSI tumors.[2] Under normal conditions the DNA mismatch repair system (MMR) ensures high fidelity DNA replication by preventing insertions and deletions (indels) of bases caused by replication slippages.[3] Replication slippage commonly occurs within microsatellites, which are small stretches of repetitive units varying from 1-6 nucleotides scattered throughout the genome.[4] A defective MMR-system, which is the underlying cause of MSI, will increase the possibility of indels in microsatellites as these are not repaired successfully. If these microsatellites are located within the coding region of a gene this will lead to frameshift mutations,[5] and since many genes have microsatellites within the coding region, MSI is often referred to as the mutator phenotype.[4] There are two subgroups within MSI-tumors, those with a low and those with a high degree of instability. MSI-L tumors are usually acknowledged to have a phenotype analogous to microsatellite stable tumors (MSS).[6]

The initial cause of MSI in CRC is most often promoter hypermethylation of the DNA mismatch repair gene *MLH1*.[7] It is not MSI *per se* which is the crucial aspect in tumorigenesis, but the mutations it generates in genes that may be involved in important regulatory signaling pathways. Because of the high background of genetic instability in MSI-

tumors it is hard to establish which of the mutated genes are playing a part in tumorigenesis. A scheme of dividing the genes into driver and by-stander genes based on mutation frequency has been suggested, and genes with mutation frequencies below 12% are considered to have no functional consequence on tumor development and be termed by-stander genes.[8]

CRC is one of the most common cancer types and has a relatively poor 5-year survival of ~60% (SEER Cancer Statistics Review, 1975-2005[*]). The most important factor for survival is tumor stage at the time of diagnosis, i.e. depth of tumor infiltration in the bowel wall, spread to lymphatic nodes and distant spread. It has been shown that MSI tumors have a better prognosis than microsatellite stable (MSS) tumors.[9] Still, a subset of patients with MSI-tumors have unfavorable prognosis, and identification of these patients might give the option of improved treatment.

In order to identify markers with prognostic value we have analyzed 41 genes, selected from a literature survey,[10] in two different colorectal tumor series (in total 202 samples).

MATERIAL AND METHODS

Hypothesis generating test series

94 sporadic, fresh frozen, colorectal carcinomas from Scandinavian hospitals had previously been selected, being classified as MSI. The Norwegian series of 42 samples was from an

[*] http://seer.cancer.gov/statfacts/html/colorect.html

unselected series of primary tumors collected between 1987 and 1989 in the South-East region of Norway.[11] Forty-six Swedish samples and six Danish samples were acquired via collaborators. These samples include no clinical data and are included to analyze mutation frequencies of the genes included. All clinico-pathological associations are therefore performed on the Norwegian tumors.

Hypothesis testing validation series

An unselected prospective series of ~950 colorectal tumor samples were collected between 1993 and 2003 at Aker University Hospital in Oslo. Samples were formalin-fixed and embedded in paraffin. All tumors were subject to a formal resection and extensive clinical data for each patient was kept. MSI status was determined for all right-sided and one third of the left-sided colonic tumors, yielding 108 MSI-positive tumors which were included in the study. No rectal tumors were included.

Determination of MSI

MSI status was determined using the consensus guidelines given by the National Cancer Institute (BAT25, BAT26, D2S123, D5S346 and D17S250).[6] The Bethesda marker protocol had to be optimized both for fresh-frozen tissue and for DNA from formalin-fixed samples. The fresh frozen samples were analyzed in a pentaplex PCR using 37ng DNA template in a 10µl reaction volume consisting of 1 x Multiplex PCR Mastermix (containing buffer, 1.5mM $MgCl_2$, nucleotides, and enzyme, QIAGEN GmbH, Hilden, Germany), primers and water (see Supplementary table for primer and PCR details), while the paraffin

embedded carcinomas was analyzed in two different PCRs: the mononucleotide markers (BAT25 and BAT26) and the dinucleotide markers (D2S123, D5S346 and D17S250) were analyzed separately using the same PCR conditions with the exception of number of cycles (mononucleotide markers – 30 cycles; dinucleotide markers – 35 cycles). Samples which showed instability at more than 20% of the markers were denoted MSI-H, the ones with only one instable marker were denoted MSI-L, while samples with no instable markers were denoted MSS.

Mutation screening

Mutation analysis of the above selected target genes was performed by fragment analysis of the microsatellite-containing regions of each gene. With the exception of *PTEN* and *EP300*, which both had repeats of interest in two separate exons, only one fragment was investigated per gene, yielding a total of 43 fragments, ranging from 61-236 bp in size. The fragments were amplified in multiplex PCRs averaging in five genes per reaction using the same PCR conditions as the pentaplex MSI-markers (see Supplementary table for primer and PCR details). When possible, the primer sequences were those that had been used in previous studies – by now some gene fragments have canonical primer sequences – the remainder were designed for this study using the Primer3 program.† Default settings were used except for adjusting melting temperatures upon occasion. All primer pairs were assessed for

† http://frodo.wi.mit.edu/cgi-bin/primer3

specificity, *i.e.* that they only amplify unique sequences of the human genome, by *in silico* PCR,[‡] and for hairpin and/or primer-dimer formation at NetPrimer.[§]

The fragments were labeled with the G5 dye set from Applied Biosystems; PET – red, NED – yellow, VIC – green, and 6-FAM – blue. The size standard used was GS500 LIZ (orange). All fragments were analyzed on a 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) using default Microsatellite Analysis settings in the GeneMapper3.7 software.

Electropherograms were visually examined for insertions/deletions by three independent authors, TA, ECR and MAM, against corresponding fragments from DNA from four different disease-free individuals. All assays were duplicated in tandem runs using different PCR machines to ensure the robustness of the results.

Mutation verification by sequencing

In order to determine the false discovery rate of the fragment analysis results, 18 of the genes (*ACVR2A*, *AIM2*, *ASTE1*, AXIN2, *BLM*, *EPHB2*, *GRK4*, *MBD4*, *PTHLH*, RAD50, *RBBP8*, *RCC2*, *SEMG1*, *SLC23A2*, *SYCP1*, *TAF1B*, *WISP3* and *ZMYND8*) were sequenced. Genomic DNA from a total of 107 samples was re-amplified with new primers including M13-tails, and directly sequenced in both directions on a 3730 Genetic Analyzer (Applied Biosystems). An initial PCR was performed using HotStar HiFidelity Polymerase (QIAGEN) following manufacture's instructions in a 25µl reaction before enzymatic clean-

[‡] http://genome.ucsc.edu/cgi-bin/hgpcr

[§] http://www.premierbiosoft.com/netprimer

up with ExoSAP-IT (USB Corporation, Ohio, USA) and sequencing PCR using BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) following the producer's instructions. Forward and reverse sequences were analyzed and compared with the genomic reference sequences. For primer details please see supplementary table.

Selection of target genes

A literature search was performed in order to survey known putative target genes, as already reported.[10] From the 162 genes identified we applied different criteria in order to maximize the genes with a possible impact on tumorigenesis which in turn will have a higher probability of determining patient survival.

For the initial selection a minimum cut-off mutation frequency of 15% across all tumor types was chosen following the assessed background mutation level 12-13%.[8;12] To as far as possible ensure that the given mutation frequency was representative, this criterion was coupled to a requirement of the existence of at least two studies of minimum twenty samples each, and a minimum of one hundred samples across all studies. These criteria returned 23 genes.

In order to include additional molecular markers a second set was defined. This included a minimum observed mutation frequency for the gene of 30%, a minimum of one study of twenty tumor samples, and involvement in one of the following categories, which were in part based on Hanahan and Weinberg's acquired capabilities of cancer: DNA repair, cell signaling, apoptosis, cell cycle, transcription or angiogenesis.[13]

On the assumption that DNA repair and cell cycle genes significantly influence the development and prognosis of MSI tumors, genes falling into these categories were sorted according to mutation frequency, and those displaying mutations in over 15% of samples were included regardless of sample number. Finally, genes that are considered by Woerner *et al.* to be true target genes in MSI CRC were included,[14] resulting in a total of 41 genes.

Statistics

2 x 2 contingency tables were analyzed using Fisher's exact test. 3 x 2 tables were analyzed by the Pearson χ^2 test. Due to a low number of samples, Mann-Whitney Wilcoxon test was performed to evaluate the difference in total number of mutations in the right-sided and rectal tumors. An independent T-test was performed when comparing continuous normally distributed data between two groups. Kruskal-Wallis one-way analysis of variance was used to test the equality of mutation frequencies among the test and validation series as well as the literature.

Five year disease-free survival was analyzed using Kaplan-Meier plots, and the Breslow test was used to compare the equality of the survival functions. Cox regression for multivariate analyses was used to determine the parameters with the greatest impact on survival. *P*-values form the likelihood ratio test was used. Disease-free survival was defined as the time from diagnose to the first event of either death from disease, locally recurrent disease or distant metastasis, while death from other reasons and death from surgery were censored. All *P* values were derived from statistical tests using the SPSS 15.0 software (SPSS, Chicago, IL, USA), and considered statistically significant at $P \leq 0.05$.

RESULTS

MSI-analysis

Tumors with known MSI-status were selected for inclusion in the test set of carcinomas. To ensure a uniform MSI threshold both in the test- and the validation set, all samples were re-analyzed using consensus markers. After re-analysis, 87 tumors were included as MSI-H and 7 as MSI-L. For the validation set, MSI analysis was performed on a total of 491 tumors (385 right-sided and 106 left-sided primary tumors). Twenty-six percent of the right-sided tumors displayed MSI-H (n = 102), 6% were MSI-L (n = 24), 61% were MSS (n = 233) and 7% were unsuccessful (n = 26). Among the left-sided tumors, 6% were MSI-H (n = 6), 9% were MSI-L (n = 10), 79% MSS (n = 84) and 6% unsuccessful (n = 6). In total, 108 (22%) of the tumors displayed a high degree of microsatellite instability, and were included in further analyses.

Mutation frequencies

Mutation frequencies of the 41 genes varied from 6% (*ZMYND8*) to 91% (*ACVR2A*) with a median of 17 (range 0-28) mutated genes per sample for the MSI-H tumors in the test series, and one (range 0-14) mutated gene per sample for the MSI-L tumors.

In the validation series, mutation frequencies for each gene varied from 4% (*EP300*) to 92% (*ACVR2A*) with a median of 19 (range 0-29) mutated genes per sample. Gene specific mutation frequencies in the two tumor series can be seen in Figure 1, in which results summarized from the literature (from [10]) are included as well.

In order to ensure that gene mutation frequencies were not simply a function of the length of the microsatellite repeats, we plotted the mutation frequencies according to increasing repeat length (Supplementary Figure). The majority of the genes (78%) have an (N)8-10 repeat, and no significant increase in mutation frequency was seen. The 5 genes with an (N)11 repeat (*ASTE1*, *MARCKS*, *MRE11A*, *PTHLH* and *TAF1B*) were significantly more frequently mutated compared with the remaining genes. As a whole, mutation rate is not a mere consequence of repeat size, exemplified by *ACVR2A*, who has one of the highest mutation frequencies (91% and 92%) with only an (A)8 repeat.

The overall mutation frequencies in the two tumor series and the literature were not significantly different from each other, as shown using Kruskal-Wallis one-way analysis of variance ($P = 0.54$). However, differences for individual genes were seen both between our series and the literature, as well as between the two tumor series. *ACVR2A*, *OGT* and *RAD50* were more frequently mutated in our analyses as compared to the literature, while *ADCY2*, *EP300*, *PA2G4* and *SEC63* were less frequently mutated. *EPHB2*, *MARCKS*, *PCNXL2*, *RBBP8*, *RCC2*, *SEMG1* and *SPINK5* were more frequently mutated in the validation series compared to the test set.

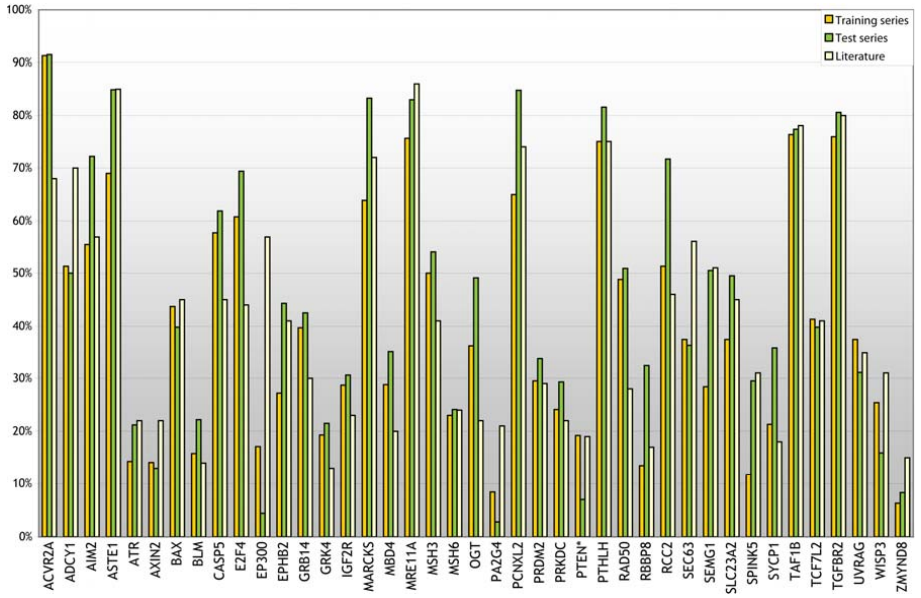


Figure 1. Mutation frequencies across two tumor series and a literature survey. MSI-H tumors from all locations (right-sided, left-sided and rectum samples in the test series (sample range: 71-87); right-sided and left-sided in the validation series (sample range: 46-108)) are included. The numbers from literature are taken from [10] and includes tumors reported as MSI and might also include cell lines and hereditary non-polyposis colorectal cancer tumors in studies not specifying this. *Only one of two exons (exon 8) was analyzed successfully in the validation series.

Seven MSI-L tumors were included in the test series to determine the effect of a suboptimal function of the mismatch repair system on microsatellite mutagenesis. Both on the individual gene level (Figure 2) and at the total level the MSI-L tumors carried significantly lower numbers of mutations compared to the MSI-H tumors. As only 7 MSI-L tumors are included, no statistics were performed.

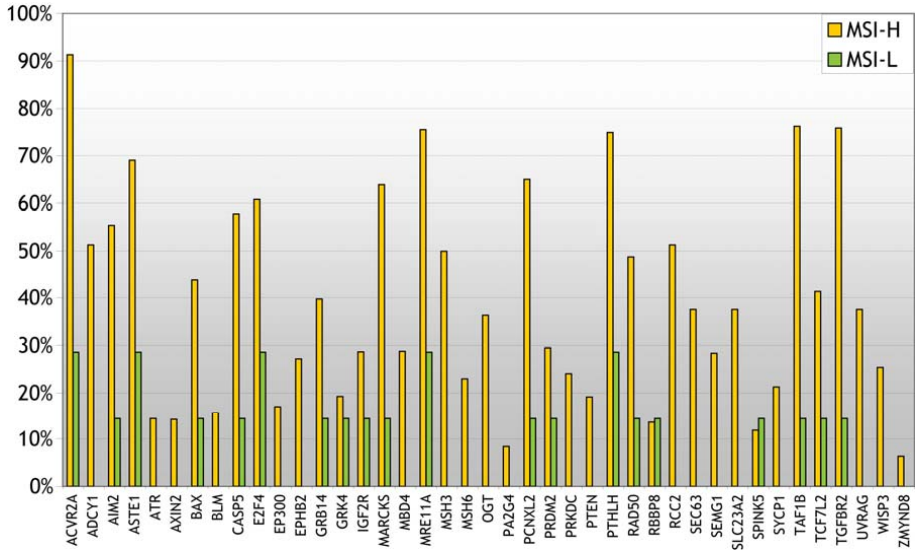


Figure 2. Mutation frequency differences between MSI-H and MSI-L tumors. Tumors with a low degree of microsatellite instability (1 out of 5 consensus markers display instability) are less prone to acquire indels within repetitive sequences.

Mutation verification by sequencing confirmed 95% of the mutations described with fragment analysis. Five of the 107 sequences differed from the fragment analysis results, yielding a false discovery rate of 5%. Figure 3 illustrates the comparisons of fragment analysis and dye-nucleotide sequencing.

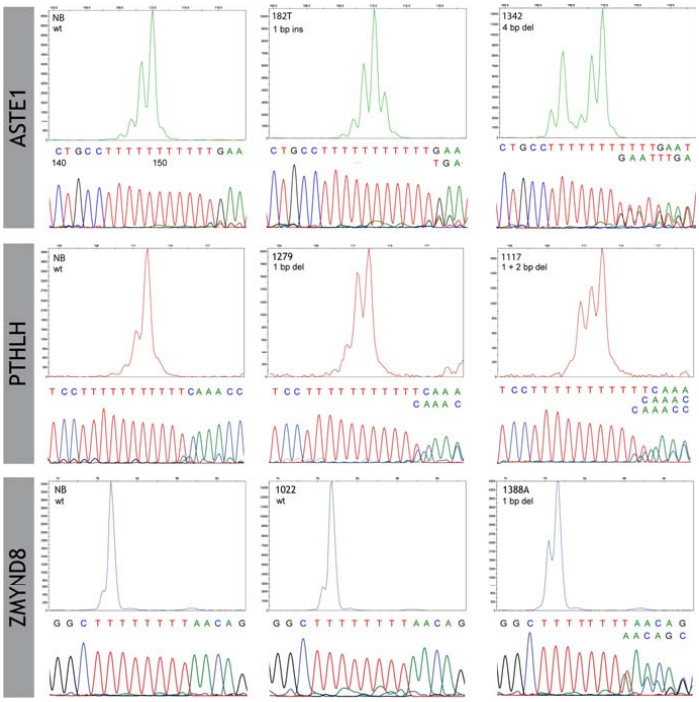


Figure 3. Representative electropherogram results and the sequence confirmation. Sample ID and mutation status as determined with fragment analysis are shown in the top left corner of each sample. The electro-pherograms are shown in the top panel for each of the *ASTE1*, *PTHLH* and *ZMYND8*. The bottom panel of each gene includes the sequence confirming the fragment analysis. The correct mononucleotide repeat is (A)11, (A)11 and (A)8, respectively. As the sequences are obtained using the reverse primer, the bases are inversed and complementary to the reference sequence.

Correlation analyses

Bimodal Pearson's correlation analyses were performed in order to detect any co-occurrence of mutations between the different genes. Mutations in *ACVR2A* and *TGFBR2* (correlation coefficient 0.54), *TAF1B* and *ASTE1* (0.57), *TAF1B* and *ACVR2A* (0.55), and *MRE11A* and *ASTE1* (0.51) correlated significantly in the test series (P = $2.9x10^{-7}$, $2.5x10^{-8}$, $1.2x10^{-7}$, $2.3x10^{-6}$, respectively).

Significant correlation of mutations between *TAF1B* and *ACVR2A* (correlation coefficient 0.58) and *MRE11A* and *ASTE1* (0.61) (P = $5.5x10^{-9}$and $4.2x10^{-10}$, respectively) were verified in the validation series. In addition, mutations in *MARCKS* and *ACVR2A* (0.57), *MRE11A*

and *TAF1B* (0.50), *ACVR2A* and *ASTE1* (0.54), *PCNXL2* and *ACVR2A* (0.51), and *MRE11A* and *PTHLH* (0.59) showed correlation greater than 0.50 (P = $1.7x10^{-7}$, $1.6x10^{-6}$, $2.9x10^{-9}$, $1.1x10^{-7}$ and $2.6x10^{-9}$, respectively).

Genetic and clinico-pathological associations

No associations were seen between individual mutation status and tumor stage or gender, except for *MSH3* and female gender (P = 0.025) in the test series. Nor when comparing the mean number of mutated genes per sample to gender or tumor stage any associations were seen.

Several of the genes show a significant difference in mutation frequency when comparing tumor location, as right-sided MSI tumors often show a higher mutation frequency than left-sided and especially rectal tumors. *ACVR2A*, *ASTE1*, *CASP5*, *MARCKS*, *MBD4*, *MRE11A*, *MSH3*, *TAF1B* and *TFGBR2* were all significantly more mutated in right-sided tumors compared to both left-sided and rectal tumors (Table 1). In addition, *PRDM2*, *BAX* and *E2F4* showed the same tendency, although not reaching a 5% level of significance (P = 0.081, 0.094 and 0.077, respectively). When comparing the mean number of mutations, the right-sided tumors carried a significantly higher number compared to the rectal tumors (mean rank 20.4 and 10.6, respectively, P = 0.008). The association between mutation frequency and site for the five MSI consensus genes, *BAX*, *IGF2R*, *MSH3*, *MSH6* and *TGFBR2*, are included in another study (Teixeira *et al.*, unpublished).

	ACVR2	*ASTE1*	*CASP5*	*MARCKS*	*MBD4*	*MRE11A*	*MSH3*	*TAF1B*	*TGFBR2*
Right colon (n = 22)	100 %	82 %	77 %	77 %	59 %	91 %	73 %	91 %	96 %
Left colon (n = 7)	71 %	57 %	43 %	29 %	0 %	57 %	57 %	57 %	57 %
Rectum (n = 7)	57 %	29 %	29 %	29 %	14 %	29 %	0 %	29 %	29 %
P -value	0.008	0.028	0.038	0.015	0.007	0.004	0.003	0.004	< 0.001

Table 1. Mutation frequencies at different tumor sites. All 9 genes with significantly different mutation frequencies in MSI-H tumors from different sites in the colon are listed.

Survival analyses

A univariate five year disease-free survival analysis of the mutation status of the 41 genes in the test series indicated that mutations in *AXIN2* and *EP300* were significantly associated with a worse prognosis while mutations in *MRE11A*, *OGT* and *RCC2* were associated with a beneficial prognosis (Table 2 top left). The positive association between survival and mutation in *RCC2* was confirmed in the validation series (Figure 4 and Table 2 bottom left). In addition, mutation in *GRK4* and *RBBP8* was associated with poor survival in the validation series (Table 2 bottom).

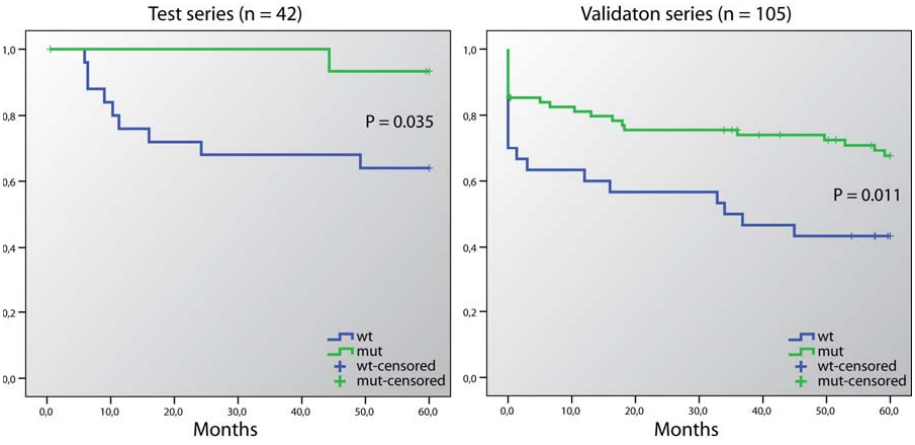


Figure 4. Survival and *RCC2* mutation. Both in the test series (left plot) and the validation series (right plot) show a significant association between mutations in *RCC2* and improved survival.

Multivariate analyses

A multivariate Cox proportional hazards model was created for both tumor series, based on all variables with a *P*-value better than 0.05 from the univariate analyses, as well as significant clinico-pathological variables (Table 2 top right). For the test series, low tumor stage at time of diagnosis was the strongest predictor of disease-free survival, followed by mutated *RCC2*, *AXIN2* wild-type and *OGT* mutation. Mutation status of *EP300* and *MRE11A* did not provide additional information. For *OGT*, no hazard ratio and confidence interval was calculated from the multivariate analysis as it lacks events in the mutation group. Advanced tumor stage and wild-type *RCC2* remained the best identifiers of poor survival in the validation series (Table 2 bottom right). In addition, wild-type *GRK4* is also associated with improved survival in the analysis (P = 0.029). The association between mutated *RCC2* and improved survival was evident only in stage I-II tumors and not in stage III-IV tumors (P = 0.004)

		Univariate			Multivariate		
	Parameter	**Survival % (SD)**	**Number of cases**	***P*-value**	**HR**	**95% CI**	***P*-value**
Test series	AXIN2						
	- wild-type	79 (6.9)	36	0.038	4.4	1.0 - 18.4	0.067
	- mutation	50 (20.4)	6				
	EP300						
	- wild-type	86 (6.6)	30	0.007			
	- mutation	50 (14.4)	12				
	MRE11A						
	- mutation	88 (6.5)	27	0.013			
	- wild-type	53 (12.9)	15				
	OGT						
	- mutation	100 (-)	15	0.011			0.008
	- wild-type	62 (9.5)	27				
	RCC2						
	- mutation	93 (6.4)	17	0.035	10.9	1.3 - 91.3	0.009
	- wild-type	64 (9.6)	25				
	Tumor stage						
	- Stage I-II	92 (5.5)	25	0.003	10.0	2.0 - 49.2	0.003
	- Stage III-IV	53 (12.1)	19				
Validation series	GRK4						
	- wild-type	66 (5.3)	83	0.019	2.2	1.1 - 4.4	0.029
	- mutation	41 (10.6)	23				
	RBBP8						
	- wild-type	65 (5.8)	72	0.024			
	- mutation	50 (8.7)	35				
	RCC2						
	- mutation	68 (5.6)	75	0.011	2.2	1.1 - 4.2	0.053
	- wild-type	43 (9.0)	30				
	Tumor stage						
	- Stage I-II	75 (5.2)	73	$1.8x10^{-6}$	3.5	1.8 - 6.8	$5.6x10^{-6}$
	- Stage III-IV	27 (8.4)	34				

Table 2. Prognostic factors for 5-year disease-free survival. Abbreviations: SD, standard deviation; HR, hazard ratio for death from disease, locally recurrent disease or distant metastasis; CI, confidence interval

DISCUSSION

Colorectal cancer is a very common disease with a poor 5-year survival. Any finding which aid in discriminating patients with a good versus poor prognosis will therefore have an impact on a large number of patients. Currently, tumor stage at time of diagnosis is the only parameter used in classifying patients into different categories, in which fit patients with stage III disease who will receive adjuvant chemotherapy.[15;16] Still, there are some patients with a stage II disease who will experience recurrence, and who would benefit from chemotherapy. On the other hand, some 50% of stage III patients will not recur after surgery alone and could be spared from the treatment related side-effects. With the aim of finding markers with prognostic potential in a low-risk group of colorectal cancer patients, we analyzed 41 genes with coding oligonucleotide repeats (all but one were mononucleotide) in two independent series of colorectal MSI tumors.

This same approach has been performed in earlier studies, but in very small numbers of both genes and patients. Results are differing, some find mutations in both of *TGFBR2* and *BAX* to be associated with poor prognosis,[17;18] another study finds them to be associated with an improved survival,[19;20] while some do not find any associations to survival at all.[21] A small study found that mutation in *ATR* showed a trend towards improved survival, although not significant.[22] A couple of studies have looked at protein expression of some of the genes included here in association to survival and found that low BAX expression is associated with poor survival,[23] and that strong staining of RAD50/MRE11/NBS1 was associated with a favorable survival.[24] In order to include driver-genes that are likely to have an impact on tumorigenesis in the study, strict selection criteria were employed.

When comparing our findings of mutation frequencies with literature, we saw that *ACVR2A*, *OGT* and *RAD50* had a higher mutation rate in our analyses, while *ADCY2*, *EP300*, *PA2G4* and *SEC63* were less frequently mutated as compared to the literature. Most of these genes have been the subject of few or single studies, often with a very restricted sample size, and therefore discrepancies were not unexpected. In the instance of *ACVR2A* another study has found a similarly high mutation frequency.[25] To the best of our knowledge, this is the first study to analyze indels in *EP300* in primary tumors. Previously it has been shown to be mutated in 4 of 7 CRC cell lines.[26] The low mutation frequency seen here indicates that it is not among the driving forces in colorectal tumorigenesis.

Some genes showed more than a 10% difference in mutation frequency between the two tumor series (*AIM2*, *ASTE1*, *EP300*, *EPHB2*, *MARCKS*, *PCNXL2*, *RBBP8*, *RCC2*, *SEMG1*, *SPINK5* and *SYCP1*). There may be biological as well as technological explanations for this discrepancy. Firstly, some biological variation is expected. Also, the fact that the test series includes a higher number of non-right-sided tumors makes us expect the mutation frequencies to be slightly lower. In fact, when comparing only right-sided tumors the mutation frequencies in the two series were more similar as only *SPINK5* and *EP300* were significantly different. All of the differently mutated genes were more often mutated in the formalin fixed validation series. It may be that the tissue fixation protocol plays a part in the elevated mutation frequencies.

Correlation analysis indicated that *ACVR2A*:*TGFBR2*, *TAF1B*:*ASTE1*, *TAF1B*:*ACVR2A*, and *MRE11A*:*ASTE1* were correlated in the test series, of which the pairs of *TAF1B*: *ACVR2A* and *MRE11A*:*ASTE1* were confirmed in the validation series. All these genes

had mutation frequencies of 75% or higher in both series. TAF1B is an essential part of the RNA polymerase I cofactor SL1, which is important for transcriptional initiation in general, as it is involved in the binding of regulatory proteins at the TATA-box.[27;28] Its co-mutated gene, *ACVR2A* (activin receptor 2), transmits the growth inhibitory effects of activin via phosphorylation of SMAD proteins to affect gene transcription through the TGFβ signaling pathway.[29] Activin has been shown to regulate differentiation, proliferation and apoptosis in several cancer types.[29] When both these genes are inactivated in a tumor, it provides severe deregulation of transcription as a whole as well as loss of the growth inhibitory effects of activin,[29] strategies which are positive for a tumor cell.

MRE11A is involved in double strand DNA break repair along with several other proteins and is clearly important in maintaining high DNA fidelity.[30] The role of *ASTE1* (previously known as *HT001*) is not known, so any mutational consequences of this gene remain to be resolved. Hence, the impact of the loss of both these genes in a tumor cell thus remains elusive. Another possibility is that the recognized correlations are merely due to high mutation rates rather than important biological functions.

MSI-L tumors are usually acknowledged to have a phenotype analogous to MSS.[6;31] As expected, we saw a significant difference in mutation frequencies between tumors with a high and low degree of microsatellite instability as the MSI-L tumors harbored a mean of one mutation per tumor. This indicates that MSI-L is insufficient to induce the mutator phenotype in CRC.

The test series included a low number of MSI-H rectal tumors which differed from the traditional right-sided MSI-tumors in mutation load, both on the individual gene mutation level as well as the total mutation level, as the rectal tumors had significantly fewer mutations. From developmental biology we know that the large intestine originates from two different embryological regions, the midgut and the hindgut, of which the midgut becomes the small intestines and proximal part of the large bowel, while the hindgut constitutes the last third of transversum and extends to the anal opening. Also, the environmental conditions in the colon differ, both the content and the passage time of the feces, making the exposure of potential mutagens different in the different regions. This difference in biology may explain the discrepancy in the downstream mutation targets. As previously mentioned, sporadic MSI-tumors are generally caused by hypermethylation of *MLH1*. In a previous study we analyzed *MLH1* hypermethylation for most of the samples included here[32]. As expected, we found that the majority of right-sided tumors had hypermethylated *MLH1* promoters. In contrast, none of the rectal tumors had this feature, indicating that MSI is caused by mechanisms other than *MLH1* methylation in these tumors. Bias from hereditary cancer, especially Lynch's syndrome has been excluded, as assessed by written questionnaires.[31]

RCC2 as a target for frameshift mutation in MSI-cancers was identified by a systematic database search in 2002 (reported using its alias *KIAA1470*).[33] However, this is the first time the mutation status of the gene is linked to survival. In this study we show that indels in a coding mononucleotide repeat in *RCC2* were significantly associated with improved survival both in the hypothesis generating and the hypothesis testing cancer series.

Previous studies have shown that RCC2 (also known as TD-60) is associated with the segregation of chromosomes in metaphase.[34] Indeed, RCC2 is critical for the integration of kinetochores into the mitotic spindle, and may be required for overall spindle assembly.[35] siRNA suppression shows that RCC2 is absolutely required for progression from prometaphase to metaphase and that its suppression activates the spindle assembly checkpoint, hence creating an effective G2/M arrest, indefinitely blocking cells from completing mitosis.[35] The same study also suggested that the presence of RCC2 is critical for the recruitment of other proteins involved in cell division to inner centromeres in mitosis, supported by a novel study which show that RCC2 is important for AURORA B (a protein involved in the chromosomal passenger complex) localization, but not activity, and that RCC2-depleted extracts were impaired in their ability to align chromosomes to the metaphase plate.[36]

These findings support the assumption that mutation in *RCC2* can be associated with improved survival as cells with this mutation will be arrested before cell division if the mutation causes transcriptional inactivation. The (A)10 repeat in *RCC2* is located in exon 1 which is in the 5' untranslated region (UTR) of the gene, 76 bases upstream of the start codon. UTRs are known to affect mRNA nuclear export, cytoplasmic localization and translational efficiency and stability.[37] The majority of translational control occurs at the level of initiation, thus implicating the 5' UTR as a major site of translational regulation.[38;39] It has also been shown that mononucleotide repeats in UTRs are conserved due to possible selective pressure relating to a functional role, and that these conserved repeats are frequently altered in MSI-cancers, and so are speculated to be involved in regulating gene expression.[40] Whether or not the indels seen in *RCC2* arrest the cells

entering M-phase remains to be seen, but the aforementioned studies combined with the positive survival data seen here indicates that this might be the case.

To summarize, by use of two independent tumor series we have showed that indels in *RCC2* divide the MSI subgroup into those with good and poor prognosis, and that this holds even after stratifying the tumors according to tumor stage.

REFERENCES

1. Grady WM, Carethers JM (2008) **Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis.** Gastroenterology 135(4): 1079-1099

2. Issa JP (2004) **CpG island methylator phenotype in cancer.** Nat Rev Cancer 4(12): 988-993

3. Kunkel TA (1993) **Nucleotide repeats. Slippery DNA and diseases.** Nature 365(6443): 207-208

4. Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) **Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review.** Mol Ecol 11(12): 2453-2465

5. Shibata D, Peinado MA, Ionov Y, Malkhosyan S, Perucho M (1994) **Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation.** Nat Genet 6(3): 273-281

6. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S (1998) **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 58(22): 5248-5257

7. Kuismanen SA, Holmberg MT, Salovaara R, de la CA, Peltomäki P (2000) **Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers.** Am J Pathol 156(5): 1773-1779

8. Duval A, Rolland S, Compoint A, Tubacher E, Iacopetta B, Thomas G, Hamelin R (2001) **Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers.** Hum Mol Genet 10(5): 513-518

9. Popat S, Hubner R, Houlston RS (2005) **Systematic review of microsatellite instability and colorectal cancer prognosis.** J Clin Oncol 23(3): 609-618

10. Røyrvik EC, Ahlquist T, Rognes T, Lothe RA (2007) **Slip slidin' away: a duodecennial review of targeted genes in mismatch repair deficient colorectal cancer.** Crit Rev Oncog 13(3): 229-257

11. Meling GI, Lothe RA, Børresen AL, Hauge S, Graue C, Clausen OP, Rognum TO (1991) **Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis.** Br J Cancer 64(3): 475-480

12. Duval A, Reperant M, Hamelin R (2002) **Comparative analysis of mutation frequency of coding and non coding short mononucleotide repeats in mismatch repair deficient colorectal cancers.** Oncogene 21(52): 8062-8066

13. Hanahan D, Weinberg RA (2000) **The hallmarks of cancer.** Cell 100(1): 57-70

14. Woerner SM, Benner A, Sutter C, Schiller M, Yuan YP, Keller G, Bork P, Doeberitz MK, Gebert JF (2003) **Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes.** Oncogene 22(15): 2226-2235

15. Duffy MJ, van DA, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, Lamerz R, Peltomaki P, Sturgeon C, Topolcan O (2007) **Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use.** Eur J Cancer 43(9): 1348-1360

16. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC, Jr. (2006) **ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer.** J Clin Oncol 24(33): 5313-5327

17. Fernandez-Peralta AM, Nejda N, Oliart S, Medina V, Azcoita MM, Gonzalez-Aguilera JJ (2005) **Significance of mutations in TGFBR2 and BAX in neoplastic progression and patient outcome in sporadic colorectal tumors with high-frequency microsatellite instability.** Cancer Genet Cytogenet 157(1): 18-24

18. Ionov Y, Yamamoto H, Krajewski S, Reed JC, Perucho M (2000) **Mutational inactivation of the proapoptotic gene BAX confers selective advantage during tumor clonal evolution.** Proc Natl Acad Sci U S A 97(20): 10872-10877

19. Jung B, Smith EJ, Doctolero RT, Gervaz P, Alonso JC, Miyai K, Keku T, Sandler RS, Carethers JM (2006) **Influence of target gene mutations on survival, stage and histology in sporadic microsatellite unstable colon cancers.** Int J Cancer 118(10): 2509-2513

20. Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, Benson AB, III, Hamilton SR (2001) **Molecular predictors of survival after adjuvant chemotherapy for colon cancer.** N Engl J Med 344(16): 1196-1206

21. Samowitz WS, Curtin K, Neuhausen S, Schaffer D, Slattery ML (2002) **Prognostic implications of BAX and TGFBRII mutations in colon cancers with microsatellite instability.** Genes Chromosomes Cancer 35(4): 368-371

22. Lewis KA, Bakkum-Gamez J, Loewen R, French AJ, Thibodeau SN, Cliby WA (2007) **Mutations in the ataxia telangiectasia and rad3-related-checkpoint kinase 1 DNA damage response axis in colon cancers.** Genes Chromosomes Cancer 46(12): 1061-1068

23. Nehls O, Okech T, Hsieh CJ, Enzinger T, Sarbia M, Borchard F, Gruenagel HH, Gaco V, Hass HG, Arkenau HT, Hartmann JT, Porschen R, Gregor M, Klump B (2007)

Studies on p53, BAX and Bcl-2 protein expression and microsatellite instability in stage III (UICC) colon cancer treated by adjuvant chemotherapy: major prognostic impact of proapoptotic BAX. Br J Cancer 96(9): 1409-1418

24. Gao J, Zhang H, Arbman G, Sun XF (2008) **RAD50/MRE11/NBS1 proteins in relation to tumour development and prognosis in patients with microsatellite stable colorectal cancer.** Histol Histopathol 23(12): 1495-1502

25. Jung B, Doctolero RT, Tajima A, Nguyen AK, Keku T, Sandler RS, Carethers JM (2004) **Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers.** Gastroenterology 126(3): 654-659

26. Ionov Y, Matsui S, Cowell JK (2004) **A role for p300/CREB binding protein genes in promoting cancer progression in colon cancer cell lines with microsatellite instability.** Proc Natl Acad Sci U S A 101(5): 1273-1278

27. Comai L, Zomerdijk JC, Beckmann H, Zhou S, Admon A, Tjian R (1994) **Reconstitution of transcription factor SL1: exclusive binding of TBP by SL1 or TFIID subunits.** Science 266(5193): 1966-1972

28. Gorski JJ, Pathak S, Panov K, Kasciukovic T, Panova T, Russell J, Zomerdijk JC (2007) **A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription.** EMBO J 26(6): 1560-1568

29. Jung BH, Beck SE, Cabral J, Chau E, Cabrera BL, Fiorino A, Smith EJ, Bocanegra M, Carethers JM (2007) **Activin type 2 receptor restoration in MSI-H colon cancer suppresses growth and enhances migration with activin.** Gastroenterology 132(2): 633-644

30. Helleday T, Lo J, van G, Engelward BP (2007) **DNA double-strand break repair: from mechanistic understanding to cancer treatment.** DNA Repair (Amst) 6(7): 923-935

31. Lothe RA, Peltomäki P, Meling GI, Aaltonen LA, Nystrom-Lahti M, Pylkkanen L, Heimdal K, Andersen TI, Moller P, Rognum TO, . (1993) **Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history.** Cancer Res 53(24): 5849-5852

32. 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.** Molecular cancer. *In press*

33. Kim NG, Rhee H, Li LS, Kim H, Lee JS, Kim JH, Kim NK, Kim H (2002) **Identification of MARCKS, FLJ11383 and TAF1B as putative novel target genes in colorectal carcinomas with microsatellite instability.** Oncogene 21(33): 5081-5087

34. Martineau-Thuillier S, Andreassen PR, Margolis RL (1998) **Colocalization of TD-60 and INCENP throughout G2 and mitosis: evidence for their possible interaction in signalling cytokinesis.** Chromosoma 107(6-7): 461-470

35. Mollinari C, Reynaud C, Martineau-Thuillier S, Monier S, Kieffer S, Garin J, Andreassen PR, Boulet A, Goud B, Kleman JP, Margolis RL (2003) **The mammalian passenger protein TD-60 is an RCC1 family member with an essential role in prometaphase to metaphase progression.** Dev Cell 5(2): 295-307

36. Rosasco-Nitcher SE, Lan W, Khorasanizadeh S, Stukenberg PT (2008) **Centromeric Aurora-B activation requires TD-60, microtubules, and substrate priming phosphorylation.** Science 319(5862): 469-472

37. Hughes TA (2006) **Regulation of gene expression by alternative untranslated regions.** Trends Genet 22(3): 119-122

38. Pickering BM, Willis AE (2005) **The implications of structured 5' untranslated regions on translation and disease.** Semin Cell Dev Biol 16(1): 39-47

39. Wilkie GS, Dickson KS, Gray NK (2003) **Regulation of mRNA translation by 5'- and 3'-UTR-binding factors.** Trends Biochem Sci 28(4): 182-188

40. Suraweera N, Iacopetta B, Duval A, Compoint A, Tubacher E, Hamelin R (2001) **Conservation of mononucleotide repeats within 3' and 5' untranslated regions and their instability in MSI-H colorectal cancer.** Oncogene 20(51): 7472-7477

Gene name (HGNC)	Multiplex Group	Annealing temperature	Number of PCR cycles	[primer] in pmol	Fluorescent dye	size in bp	Forward	Reverse
Fragment analysis primers								
BAX	1	58°C	27	0.8	VIC	94	ATCCAGGATCGAGCAGGGCG	ACTCGCTCAGCTTCTTGGTG
IGF2R	1	58°C	27	1.0	NED	111	GCAGGTCTCCTGACTCAGAA	GAAGAAGATGGCTGTGGAGC
PRKDC	1	58°C	27	1.6	6-FAM	113	GACTCATGGATGAATTTAAAATTGG	TTTGAAAATAACATGTAAATGCATCTC
SLC23A2	1	58°C	27	1.0	6-FAM	61	GACTACTACGCCTGTGCACG	TGTTTATTGCGTGGATGGG
TGFBR2	1	58°C	27	1.0	NED	73	CTTTATTCTGGAAGATGCTGC	GAAGAAAGTCTCACCAGG
AXIN2	2	60°C	27	1.6	6-FAM	123	CCTACCCCTTGGAGTCTGC	CAGGGTCCTGGGTGAACA
GRK4	2	60°C	27	1.6	NED	68	GGAAAAATGTATGCCTGCAA	GCCATAGCTTCACCTTTCCTC
MSH2	2	60°C	27	1.6	VIC	179	AGCAGTCAGAGCCCTTAACC	GCCATTTAAAGCTAGTTATCTAATCCA
PTHLH	2	60°C	27	1.6	PET	113	TTTCACTTTCAGTACAGCACTTCTG	GAAGTAACAGGGGACTCTTAAATAATG
SEC63	2	60°C	27	1.6	NED	145	GGAGGATGGCAACAGAAGAG	CATTCCCAACGACTCCATTT
SPINK5	2	60°C	27	1.2	VIC	99	TGAGGCGTTTGTTCACTTTG	TGCTCCTGTCTTCATCCTCTT
ASTE1	3	60°C	27	1.0	VIC	117	ATATGCCCCCGCTGAAATA	TTGGTGTGTGCAGTGGTTCT
ATR	3	60°C	27	2.0	PET	70	GCTTCTGTCTGCAAGCCATT	TGAAAGCAAGTTTTACTGGACTAGG
EP300, ex27	3	60°C	27	1.0	PET	156	ACACAACAGGGCATATTTGG	CATGGACAATACGCTCTGATACA
RAD50	3	60°C	27	1.0	NED	86	GCGACTTGCTCCCAGATAAAC	GCACAAGTCCCAGCATTTCA
ACVR2A	4	58°C	27	1.6	VIC	113	GTTGCCATTTGAGGAGGAAA	CAGCATGTTTCTGCCAATAATC
MBD4	4	58°C	27	2.0	NED	98	TGACCAGTGAAGAAAACAGCC	GTTTATGATGCCAGAAGTTTTTTG
OGT	4	58°C	27	1.2	PET	116	TCACTTTTGGCTGGTCAGAG	GGGAGGGAAAGGAGGTAAAG
SYCP1	4	58°C	27	1.6	6-FAM	153	CCCCTTCATCTCTAACAACCC	CACTGATTCTCTGAAATTAAACAAATAAC
TCF7L2	4	58°C	27	1.6	6-FAM	75	GCCTCTATTCACAGATAACTC	GTTCACCTTGTATGTAGCGAA
CASP5	5	60°C	27	1.2	PET	141	CAGAGTTATGTCTTAGGTGAAGG	ACCATGAAGAACATCTTTGCCCAG
EP300, ex3	5	60°C	27	1.2	VIC	100	GCCTTGGTCTCCAGATTCAG	ATGTTGGGCATTCCTCCA
EPHB2	5	60°C	27	1.2	VIC	85	CACGAGACGTCACCAAGAAA	CCCCTCCCAGGATCTGTT
PA2G4	5	60°C	27	1.2	6-FAM	93	GCAAGTCGAAAAACCCAGAA	GGGAGAAAGAGGGATCAGGT
SEMG1	5	60°C	27	1.2	NED	152	ACAACGACCGAAACCCATTA	CCCACAAAAGTCCTGGAGAG
PCNXL2	6	57°C	27	1.0	NED	121	GGAAAATTATGAACAGCCACAA	GCAGCCAAATGCTTGTTATG
RCC2	6	57°C	27	1.6	VIC	143	GCATTTGTTCTGGAAGCTCGT	GTGATGAGAAACCGGAGAGAA
TAF1B	6	57°C	27	1.2	VIC	115	CCAAATAAAAGCCCTCAACC	TGTCCTGACATCATGAAGGTG
ZMYND8	6	57°C	27	1.2	6-FAM	83	CAAAGAAGGTTGTCAGATGGAC	CCTCTTTAATCTCCACTGGG
GRB14	7	53°C	27	1.6	6-FAM	80	GGCAATAGTGATATTTATGT	GGCTTAAAGCAGAATCCAT
MRE11A	7	53°C	27	1.6	NED	122	AATATTTTGGAGGAGAATCT	AATTGAAATGTTGAGGTTGCC
PRDM2	7	53°C	27	1.2	VIC	145	TCTCACATCTGCCCTTACTG	GTGATGAGTGTCCACCTTTC
PTEN, ex7	7	53°C	27	1.6	6-FAM	236	CCTGTGAAATAATACTGGTATG	GTTTCTTCTCCCAATGAAAGTAAAGTACA
PTEN, ex8	7	53°C	27	1.6	PET	166	GTGCAGATAATGACAAGGAATA	GTTTCTTACACATCACATACATACAAGTC
AIM2	8	60°C	27	1.6	PET	76	TTCTCCATCCAGGTTATTAAGGC	TTAGACCAGTTGGCTTGAATTG
E2F4	8	60°C	27	2.4	PET	97	TGCAGTCTTCTGCCCTGCT	GGGTTGGGTCCGGACGAA
MSH6	8	60°C	27	1.2	6-FAM	94	GGGTGATGGTCCTATGTGTC	CGTAATGCAAGGATGGCGT
RBBP8	8	60°C	27	2.0	VIC	88	GAATACAAGTTTGTCCCCTTC	GCTAGATATACAAGTGTTGCTA
ADCY1	9	58°C	27	1.2	NED	98	CCAGAAGCAAATTCACAAGAC	TTTTGCGTGTTCCTTCCTTC
BLM	9	58°C	27	1.2	6-FAM	88	GAGTAGCAACTGGGCTGAAA	GACAGCAGTGCTTGTGAGAA
MSH3	9	58°C	27	1.2	VIC	153	AGATGTGAATCCCCTAATCAAGC	ACTCCCACAATGCCAATAAAAAT
MARCKS	10	58°C	30	3.5	PET	109	GACTTCTTCGCCCAAGGC	GCCGCTCAGCTTGAAAGA
UVRAG	10	58°C	30	4.5	PET	116	TTTATTTTTAAACATTGTGAGTATG	TTTTTAACTGCAGGCATTCAC
WIPS3	10	58°C	30	7.0	PET	126	TCTCCCTTTGTTTTAGC	ATTGGTCACCCTGTTAG
Consensus MSI-primers								
BAT25	NA	55°C	27	1.2	NED	124	TCGCCTCCAAGAATGTAAGT	TCTGCATTTTAACTATGGCTC
BAT26	NA	55°C	27	1.6	6-FAM	122	TGACTACTTTTGACTTCAGCC	AACCATTCAACATTTTTAACCC
D2S123	NA	55°C	27	1.6	NED	211	AAACAGGATGCCTGCCTTTA	GGACTTTCCACCTATGGGAC
D5S346	NA	55°C	27	1.2	VIC	125	ACTCACTCTAGTGATAAATCGGG	AGCAGATAAGACAGTATTACTAGTT
D17S250	NA	55°C	27	3.0	6-FAM	152	GGAAGAATCAAATAGACAAT	GCTGGCCATATATATATTTAAACC
Sequencing primers								
ACVR2A	NA	58°C	40	10.0	NA	311	CCAGTTTGAAAGTCAGGAGGA	TGTGAAGATCACCTTCCAGAAA
AIM2	NA	58°C	40	10.0	NA	348	GATCCAAGGCAGACCAATGT	TTCTGAATTTGTTTATCCAGCAA
ASTE1	NA	58°C	40	10.0	NA	395	TCCTGTTGCACTGAATTACTTCTT	AACTGAGTTTTATTCAATGTTGGAG
AXIN2	NA	58°C	40	10.0	NA	398	CCCAGTTTCTTTCCTTCTGTTTT	TTCTCATGGGAGGGTTTGAG
BLM	NA	58°C	40	10.0	NA	392	CCCTATGGAGGGTGATTCCT	CCCAGTCATCATCATCATCAA
EPHB2	NA	58°C	40	10.0	NA	322	CTCGGCTCACCTCTTCCTC	TGGACACATCGCATGAATCT
GRK4	NA	58°C	40	10.0	NA	298	CCTAAGAAATGCCAGGTGGA	AATGACTTCCACGGCTTCAG
MBD4	NA	58°C	40	10.0	NA	302	CTGCATTTCTGATGCTGGAG	TTGGTGAGCAGTTGTTGTCC
PTHLH	NA	58°C	40	10.0	NA	370	CCACGGCTAAAAGACCTGAG	GATGGGTTTGCCAGCTTAAA
RAD50	NA	58°C	40	10.0	NA	276	CCCCGTTTGTCAGAGAGTTT	GGCATACCAGCTCAGAGTCC
RBBP8	NA	58°C	40	10.0	NA	398	TTCCTCTGCTTTTCCCCTTC	TGATGTGTGAAAAGGGCACTA
RCC2	NA	58°C	40	10.0	NA	328	CCGCACATGTGTTTCTGTTT	ATCCGCCTTCCTTCCTCTT
SEMG1	NA	58°C	40	10.0	NA	246	AGTGATCGTCATTTGGCACA	GGGGAGGCTCATCTTCCTAC
SLC23A2	NA	58°C	40	10.0	NA	341	ACTGGTCCTGGTCACTTTCG	GTCGCTAGAGTCCTGCTGCT
SYCP1	NA	58°C	40	10.0	NA	228	TGTACTCAGGCCCCTTCATC	TTTGGCTCTGGCAAATAAGAA
TAF1B	NA	58°C	40	10.0	NA	266	GGTCCTGGCACTCACAATTT	TGTCCTGACATCATGAAGGTG
WISP3	NA	58°C	40	10.0	NA	350	ATTTTGGGGTTGGGAAAGAG	CTGTCGCAAGGCTGAATGTA
ZMYND8	NA	58°C	40	10.0	NA	376	CCACATGCTGGGTTTTCTTT	TTTATGGGGTGAGGTGAAGG

Supplementary Table. Primer and PCR details.

Abbreviations: HGNC - Human Genome Nomenclature Comittee; NA - not applicable; pmol - picomol; bp - base pair

Appendix I:
List of abbreviations

Appendix II:
Associated article:
Ellen C Røyrvik, Terje Ahlquist, Torbjørn Rognes and Ragnhild A Lothe.

Slip slidin' away: a duodecennial review of targeted genes in mismatch repair deficient colorectal cancer.

Critical Reviews in Oncogenesis. 2007 Dec;13(3):229-57.

Appendix I – List of abbreviations

5-FU – 5' fluorouracil

CEA – carcinoembryonic antigen

CIMP – CpG island methylator phenotype

CIN – chromosomal instability

CpG – Cytosine followed by a guanine, bound with a phosphodiester binding

CRC – colorectal cancer

DHLPC – denaturing high performance liquid chromatography

DNA – deoxyribonucleic acid

DNMT – DNA methyl transferase

FAP – familial adenomatous polyposis

FOBT – faecal occult blood test

HNPCC – hereditary non-polyposis colorectal cancer

HP – hyperplastic polyp

MAP – mitogen activated protein

MBD – methyl binding domain

MLPA – multiplex ligation-dependent probe amplification

MMR – DNA mismatch repair

MSI – microsatellite instability

MSP – methylaton-specific polymerase chain reaction

MSS – microsatellite stable

PCR – polymerase chain reaction

RNA – ribonucleic acid

RT-PCR – real time polymerase chain reaction

TEA – triethylammonium ion

WNT – Wingless-type MMTV integration site family

Critical Reviews™ in Oncogenesis, 13(3):229–257 (2007)

Slip Slidin' Away: A Duodecennial Review of Targeted Genes in Mismatch Repair Deficient Colorectal Cancer

Ellen C. Røyrvik,[1,2] Terje Ahlquist,[1,2] Torbjørn Rognes,[3,4] & Ragnhild A. Lothe[1,2,5]

[1]Department of Cancer Prevention, Institute for Cancer Research, Rikshospitalet-Radiumhospitalet Medical Center, Norway

[2]Center for Cancer Biomedicine, University of Oslo, Norway

[3]Center for Molecular Biology and Neuroscience, Rikshospitalet-Radiumhospitalet Medical Center

[4]Department of Informatics, University of Oslo, Norway

[5]Department of Molecular Biosciences, University of Oslo, Norway

Address all correspondence to Ragnhild A. Lothe, Department of Cancer Prevention, Institute for Cancer Research, Rikshospitalet-Radiumhospitalet Medical Center, Norway; rlothe@rr-research.no

ABSTRACT: Roughly 15% of colorectal tumors are characterized by microsatellite instability (MSI), a deficiency caused by defective DNA mismatch repair, which leads to profuse insertions and deletions in microsatellites. Downstream target genes of this defective repair are those prone to exhibit these insertion/deletion mutations in their coding regions and potentially having functional consequences in, and providing a growth advantage for, the cancer cell. This review presents the last 12 years of research on these MSI target genes, systematizing the mutation details of the more than 160 genes identified to date, and includes their mutation frequencies in colorectal and other MSI (e.g., gastric and endometrial) tumors. Functional aspects of certain targets and the target gene concept itself are also discussed, as is the comparative wealth of potential target genes—assessed by scanning the coding sequences of the human genome for mononucleotide repeats—yet to be investigated.

KEYWORDS: Colon carcinoma, microsatellite instability, coding microsatellites, DNA mismatch repair, target genes

I. INTRODUCTION

Colorectal cancer is one of the most common cancer types to affect both sexes.[1] It is thought to derive from stem cells in colonic crypts, where there is a naturally high cellular turnover, a characteristic that may well favor the acquisition of neo-

0893–9675/07/$35.00

plastic properties.[2] Colorectal tumors develop through waves of clonal expansion following events that confer a growth advantage upon a cell.[3-5] Some sort of pronounced genetic or epigenetic instability is usually considered to be a prerequisite for the development of tumors.[4,6] Sporadic cases of colorectal cancer can be sub-classified according to their molecular phenotypes: one referred to as chromosomally unstable (CIN), and the other unstable in microsatellites (MSI), representing 85% and 15% of all cases, respectively.[7] The former type is characterized by its widespread aneuploidy,[8] which may be a result of the instability,[9] and is largely overlapping with the category of tumors designated as microsatellite stable (MSS). The MSI-type carcinoma, the focus of this review, is usually diploid or near-diploid, but exhibits instability in microsatellites, short DNA sequences consisting of a stretch of small repetitive units (1-6 nucleotides) flanked by unique sequences. MSI tumors are also prone to lymphocyte infiltration, have a higher incidence of proximal colonic location, and have poorer differentiation than MSS tumors.[10-17] The largely proximal location of these tumors may be significant in that this area of the colon is of a different embryological provenance than the distal portion, the proximal originating from the midgut and the distal from the hindgut.[18] In addition, there are differences in blood supply, metabolism, bacterial flora, antigenic profile, and gene expression between the two sections of the normal colon.[18-20] All of the above have led to the assumption that there are at least two distinct pathways which may lead to colon cancer – each with their own set of preferential, though not absolute, molecular alterations and locations. A third pathway of colorectal tumorigenesis has been proposed, the CpG island methylator phenotype (CIMP), based on concurrent promoter hypermethylation of multiple genes in a given tumor.[21] However, the existence of this phenotype is controversial, and no consensus regarding it has been reached.

In the progression from adenoma to carcinoma, both the MSI and MSS phenotypes often display early mutations in either the CRC "gatekeeper" genes *APC* (at a frequency as high as 80%[22]) or *CTNNB1* (β-catenin), and in *KRAS* for the development of adenomas, but their genetic and epigenetic profiles diverge more significantly when they start to approach malignancy (Fig. 1).[4] Though subsequent gene targets may be disparate, mutations frequently affect the same pathways in MSI and MSS tumors, as is the case with the *TGFβRII-SMAD*s pairing involved in proliferation repression and *TP53-BAX* in apoptotic promotion (Fig. 2). Signaling pathways such as MAPK, WNT, TGFβ, AKT, and the TP53 network are affected in most CRCs, and they are all implicated in cell cycle control. Some common targets in these pathways are illustrated in Figure 2; their alterations are the result of both genetic and epigenetic mechanisms.

II. IDENTIFICATION OF THE INSTABILITY OF TUMORS

The hereditary non-polyposis colorectal cancer syndrome (HNPCC) accounts for less than 5% of all CRC. The position of a gene co-segregating with this disease

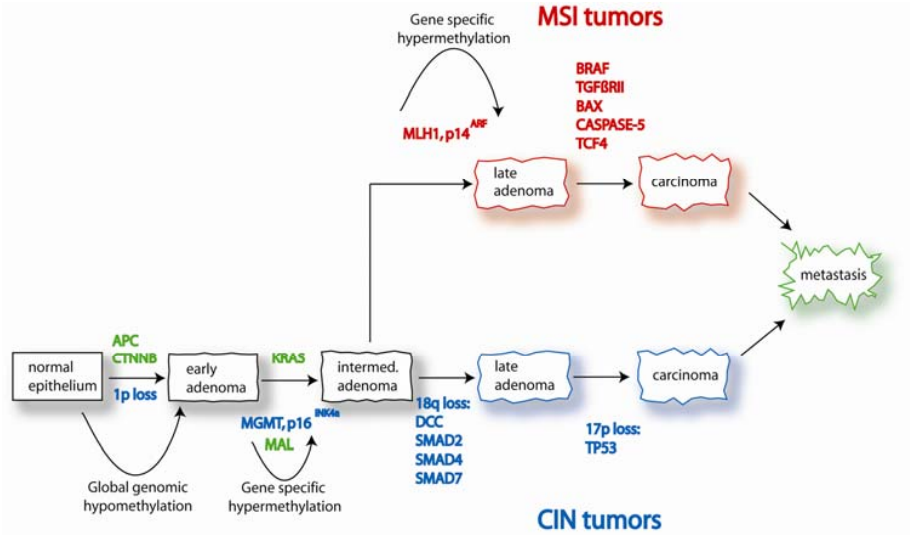


FIGURE 1. Adenoma-carcinoma sequence**.** Changes affecting CIN tumors in blue, MSI tumors in red; those common to both are in green.

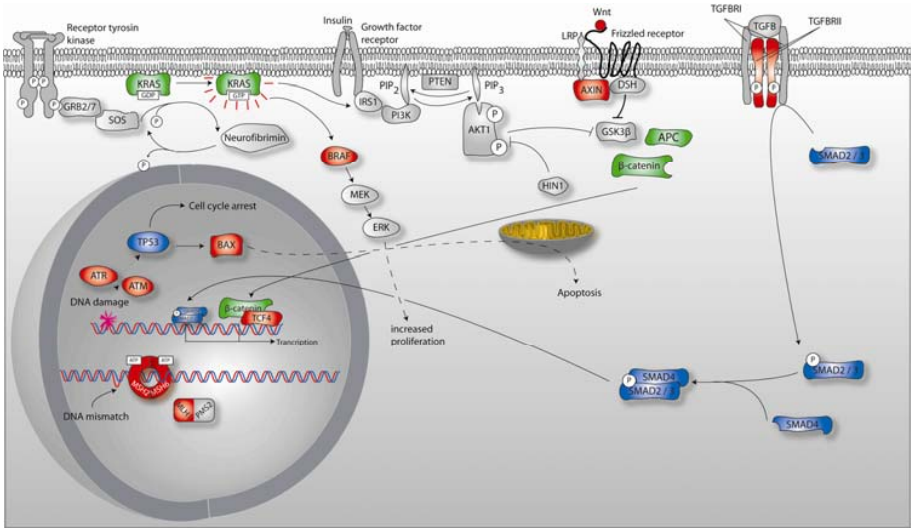


FIGURE 2. Signaling pathways commonly affected in colorectal tumorigenesis**.** Proteins in red indicate genes often affected in MSI tumors due either to mutations or epigenetic events; those in blue are genes targeted by other mutational or epigenetic mechanisms in MSS tumors.

was, by reverse genetics, found to be located at chromosome bands 2p15-16 in 1993.[23] Later that year, two groups isolated the gene in question, MSH2, which, if mutated in the germline, causes HNPCC.[24,25] Studies of HNPCC tumors revealed a ladder of novel alleles, rather than the expected loss of one allele expected according to the two-hit hypothesis for the inactivation of a potential tumor suppressor gene.[10,12-14] In this manner, colorectal carcinogenesis was tied to a novel mechanism—defective mismatch repair. A variety of tumor types are found in patients with HNPCC, the most common being colorectal, endometrial, gastric, and ovarian tumors.[26]

Tumors with MSI were initially dubbed replication error tumors (RER+), and the classification was often based on analyses of a variable number of dinucleotide loci.[10,12-14] Scoring of a tumor as RER+ or RER- was somewhat arbitrary as it depended on the total number of loci analyzed. A consensus panel of five mono- and dinucleotide markers (BAT25, BAT26—mononucleotide; D2S123, D5S346, D17S250 - dinucleotide) was implemented later, and this so-called Bethesda panel is now the current standard for assessing microsatellite instability in both the hereditary and sporadic colorectal cancers.[27] The three dinucleotide markers depend on the availability of corresponding normal DNA in order to score all affected tumors, whereas the BAT markers, being quasimonomorphic, can be more confidently used independent of non-tumor DNA. Furthermore, the BAT markers are highly informative, and therefore it has been suggested that these two, or even just BAT26, are sufficient for population-based diagnostics aiming to identify cases with potential germline cancer predisposition.[28]

It has long been suspected that patients with MSI-CRC have a better overall survival than those with MSS,[10,13,15] and it now seems well-established that the former phenotype is associated with an improved prognosis to the degree of at least 15% compared to patients with the latter type.[29] Diploidy is also a marker for positive prognosis, and though the majority of MSI-H tumors are diploid, ploidy and MSI status appear to be independent markers, as diploidy was indicative of increased survival even within the MSI group.[30]

III. DEFECTIVE MMR: THE GENERATOR OF THE MSI PHENOTYPE

The MSI phenotype is caused by faulty or lacking mismatch repair proteins of the MutL, MutS homolog repair systems, which then fail to correct insertions and deletions primarily caused by replication slippage in microsatellites with small repetitive units. Replication slippage is liable to occur at such sequences; following a transient, local dissociation of the nascent parental DNA strand, reassociation occurs between misaligned complementary repeat units, thereby lengthening or shortening the newly synthesized strand.[31-33] If the MMR system is defective, such errors will not be repaired and will accumulate in the cell. Most of the sporadic MSI tumors are caused by the silencing of *MLH1* through

promoter methylation, but some are caused by LOH/somatic mutation in *MSH2*.[34-40]

In the first step of eukaryotic mismatch repair, MSH2 recognizes the error and forms a heterodimer with either MSH3 or MSH6, depending on the nature of the irregularity. MSH3 is specific for insertion/deletion (indel) loops of 2-4 nucleotides, while MSH6 is specific for single nucleotide loops or mismatches. A complex of DNA, MutS homologs, and ATP recruits a heterodimer of MLH1 and PMS2, which displaces the main processive DNA polymerase δ and the sliding clamp PCNA, before recruiting base excision machinery to remove the tract in which the mismatch occurs. As the binding partners of MSH2 and MLH1 are completely dependent on them, any defect in MSH2, MSH3, and MSH6 will effectively inhibit the functions of PMS2 (as well as PMS1 and MLH3, of uncertain functional relevance) for MLH1.[41,42]

IV. CELLULAR CONSEQUENCES OF DEFECTIVE MMR

Defects in the systems above are not in themselves carcinogenic, they simply provide the occasion for accumulation of indels in the existing microsatellites in the genome—the resulting changes in *some* of the affected genes enables the incipient tumor to acquire necessary carcinomatous characteristics such as evasion of apoptosis, lack of dependency of extracellular/extratumoral growth signals, insensitivity to anti-growth signals, angiogenesis and unlimited scope for replication.[43] The search for such target genes of MSI-CRC first bore fruit in the shape of the tumor suppressor gene *TGFβRII* in 1995,[44,45] two years after the MSI phenotype was described for familial cancers,[10,12,14] to be followed by what were to become the other "canonical" target genes, partly by virtue of their early discovery and subsequent frequent assessment: *BAX,*[46] *IGFIIR,*[47] *MSH3*, and *MSH6*.[48] In a stepwise model of tumorigenesis, the existence of "secondary mutators" like MSH3 and MSH6[*] has been suggested, given that they are DNA repair proteins prone to frameshift mutations themselves, and may, when mutated, exacerbate the phenotype.[49,50] MRE11, through its probable MLH1-related involvement in 3' nick-directed MMR, may have a similar effect.[51]

Cancers of the hereditary non-polyposis colorectal cancer (HNPCC) type, the most common being colorectal and endometrial,[26] also display microsatellite instability (see above), and have similar mutational spectra to the sporadic MSI cancers, though the initial genetic flaws in these cases are germline mutations in the mismatch repair systems.[49]

V. LITERATURE SURVEY OF TARGET GENES

In order to survey known and putative target genes, a search was performed in

[*] This, naturally, can only be the case if the tumor possesses functional MSH2.

the PubMed database; search terms were compositions of *mononucleotide, microsatellites, repeats, genes, frameshift*, and the MeSH term *neoplasms*. From the relevant articles, with the cut-off date for inclusion being December 2006, mutation frequency data for genes with coding microsatellites (or microsatellites in introns and UTRs, as mentioned below) was pooled. Included were studies on MSI tumors and cell lines of all tissue types in which the MSI phenotype occurs, the overwhelming majority being gastrointestinal and endometrial tumors. Inclusion of mutation data on cell lines was kept to a minimum, as cell lines cannot be used to represent the situation in primary tumors with regard to the frequencies of frameshift mutations of this type.[44,49] Table I encompasses those genes and studies suitable for pooling, across tissue types but with particular attention to sporadic, primary colorectal tumors.

TABLE I

Mutation frequencies of microsatellite-containing genes

Gene symbol—the most common name for any given gene in the scientific corpus surveyed; *HGNC symbol*—the name approved by the Human Genome Nomenclature Committee, updated to October, 2007; *Repeat*—the repeat unit and number of units of microsatellites that were tested; *Gene location*—the location of the repeat within the gene, taken when possible from the scientific corpus, otherwise from Ensembl Exon View of a representative transcript; *Chr. Location*—the location of the gene on a chromosome, taken when possible from the scientific corpus, otherwise from Ensembl or GeneCards.org; *Mut%(tot)*—the mutation frequency across all studies of any microsatellite unstable tumor type; *Mut S/C*—the number of mutated samples across studies of colorectal carcinomas (mostly sporadic, but with HNPCC cases and cell lines where these could not be separated) over the total number of samples; *Mut% S/C*—the mutation frequency across studies of colorectal carcinomas (mostly sporadic, but with HNPCC cases and cell lines where these could not be separated); *Ref*—Articles from which mutation frequency information was pooled.

Gene	HGNC	Repeat	Chr. Location	Mut % (tot)	Mut. S/C	Mut %(S/C)	Refs.
ABCF1	ABCF1	(A)10	6p21.33	29 %	17/58	29 %	44,52
AC1	C4orf6	(T)10	4p16.2	68 %	14/20	70 %	53,54
ACTRII	ACVR2A	(A)8	2q22.3	70 %	95/140	68 %	55-59
AD7c-NTP		(T)8	1p36	6 %	2/35	6 %	55
AIM2	AIM2	(A)10	1q22	52 %	46/81	57 %	44,55,56
AMYB	MYBL1	(A)8	8q22	11 %			56
ANG2	ANGPT2	(A)9	8p23.1	4 %	2/57	4 %	49
APAF-1	APAF1	(A)8	12q23	8 %	5/79	6 %	60-62
ATM	ATM	(T)7	11q22-23	13 %	4/44	9 %	63-65
ATR	ATR	(A)10	3q23	23 %	55/252	22 %	17,44,49,53,66,67
AXIN2	AXIN2	(G)7, (C)6, (A)6, (C)5	17q24.1	20 %	18/81	22 %	59,68,69

TABLE I, continued
Mutation frequencies of microsatellite-containing genes

Gene	HGNC	Repeat	Chr. Location	Mut %(tot)	Mut. S/C	Mut %(S/C)	Refs.
BAT1	BAT1	(T)8	6p21.3	16 %			56
BAX	BAX	(G)8	19q13.3-q13.45	42 %	359/773	45 %	46,49,53,55,56,59,62,67,70-109
BCL10	BCL10	(A)8	1p22	8 %	14/172	8 %	60,66,95,105
BLM	BLM	(A)9	15q26.1	16 %	51/373	14 %	17,49,55,59,64,66,67,90,96,110,111
BLYM		(A)8	4q28.1	5 %	5/96	5 %	66
BRCA1	BRCA1	(A)8	17q21	2 %	3/126	2 %	17,64,71,110
BRCA2	BRCA2	(A)8	13q12.3	2 %	6/191	3 %	17,49,71,102,110
CANX	CANX	(T)8	5q35.3	21 %			56
CASP1	CASP1	(A)8	11q23	4 %	0/78	0 %	56,62,71
CASP4	CASP4		11q22.3	0 %	0/9	0 %	94
CASP5	CASP5	(A)10	11q22.2-q22.3	43 %	94/207	45 %	17,44,49,53,59,62,87,99,106,107
CBL	CBL	(ATG)6	11q.23.3	12 %	1/11	9 %	96,110
CBP*	CREBBP	(C)5	16p13.3	86 %		86 %	112
CCDC28A	CCDC28A	(A)8	6q24.1	10 %	3/41	7 %	55,56
CCKBR	CCKBR	(T)8	11p15.4-p15.5	19 %	2/15	13 %	113
CDC25C	CDC25C	(A)8	5q31.2	11 %	10/93	11 %	66
CDX2	CDX2	(G)7	13q12.2	2 %	1/81	1 %	59,62,114
CEBPZ	CEBPZ	(A)9	2p22.2	14 %	20/148	14 %	49,66,115
CHD2	CHD2	(A)10	15q26	12 %	7/58	12 %	44,52
CHK1	CHEK1	(A)9	11q24.2	9 %	9/68	13 %	84,94,102
CRSP3	MED23	(T)8	6q22.33-q24.1	3 %	1/38	3 %	105
CYSLT1	CYSLTR1	(A)8	Xq21.1	9 %	4/44	9 %	55
DD5	EDD1	(A)8	8q22	25 %			54,56
DNA-PKcs	PRKDC	(A)10	8q11.21	22 %	50/228	22 %	44,49,53,61,66,67
Doc-1	FILIP1L		3q12.1	2 %	1/57	2 %	49
DRP	INPPL1	(C)8	11q13.4	4 %	2/42	5 %	71
DSTN	DSTN	(T)8	20p12.1	12 %	4/42	10 %	55,56
E2F-4	E2F4	(CAG)13*	16q21-22	47 %	50/111	45 %	73,92,94,99,103,109,116,117
EIF5	EIF5	(CAC)7	14q32.32	0 %	0/11	0 %	110
ELAVL3	ELAVL3	(G)9	19p13.2	37 %	7/19	37 %	52,54
EP300	EP300	(A)5, (A)7	22q13.2	57 %	4/7	57 %	112
EPHB2	EPHB2	(A)9	1p36.1-p35	41 %	101/246	41 %	118
ERCC5	ERCC5	(A)9	13q33.1	9 %	8/93	9 %	61,66
F8		(A)8 * 2	Xq28	15 %	6/41	15 %	55
FACE-1	ZMPSTE24	(T)9	1p34.2	8 %	3/37	8 %	55

TABLE I, continued
Mutation frequencies of microsatellite-containing genes

Gene	HGNC	Re-peat	Chr. Loca-tion	Mut %(tot)	Mut. S/C	Mut %(S/C)	Refs.
FAS	FAS	(T)7	10q23.31	7 %	3/30	10 %	60,107
FLASH	CASP8AP2	(A)9	6q15	0 %	0/13	0 %	61
FLJ11186	C14orf106	(A)11	14q13.1-14q21.3	64 %	25/39	64 %	44
FLJ11222	MNS1	(A)10	15q11.2	28 %	11/39	27 %	44
FLJ11383	PCNXL2	(A)10	1q42.2	74 %	29/39	74 %	44
FLJ11712	RNASEH2B	(A)10	13q14.3	18 %	7/39	18 %	44
FLJ13615	CEP290	(A)11	12q21.33	28 %	11/39	28 %	44
FLJ20139		(A)10	1p21.2	31 %	12/39	31 %	44
FLT3LG	FLT3LG	(C)9	19q13.3	36 %	7/20	35 %	53,54
FTO	FTO	(T)14	16q12.2	80 %	16/20	80 %	53
GART	GART	(A)10	21q22.11	22 %	13/60	22 %	44,53
GR6	C3orf27	(GA)9	3q21.3	17 %	3/18	17 %	119
GRB-14	GRB14	(A)9	2q24.3	30 %	17/57	30 %	49
GRK4	GRK4	(A)9	4p16.3	13 %	19/148	13 %	49,66
HBP17	FGFBP1	(A)8	4p15.32	8 %	3/38	8 %	105
HDCMA18P	LARP7	(A)8	4q25	16 %	3/44	7 %	55,56
MSH2	MSH2	(A)27	2p21	63 %	22/35	63 %	15
MSH3	MSH3	(A)8	5q14.1	40 %	337/831	41 %	17,48,49,53,55,56, 59,62,66,67,70,71, 75,77,82,85,87-94, 96-104,110,113, 117, 120-123
MSH6	MSH6	(C)8	2p16.3	25 %	168/712	24 %	48,49,53,55,56,59,62, 64,67,70,71,75,77,82, 85,87-94,96,99-104,110,121-123
hnRNP	HNRPH1	(T)8	5q35.5	22 %			56
HPDMPK	FBXO46	(T)14	19q13.32	95 %	19/20	95 %	53
RAD50	RAD50	(A)9	5q23.3	32 %	42/148	28 %	17,49,59,61,64,67
HT001	ASTE1	(A)11	3q21.3	86 %	17/20	85 %	52,54
HTF34	ZNF93	(A)8	19p13.1-p12	7 %	9/124	7 %	66,105
IDN3	NIPBL	(A)8	5p13.2	7 %	3/44	7 %	55
IGF IIR	IGF2R	(G)8	6q25.3	22 %	120/530	23 %	46,47,49,53,55,59, 70,74,75,80,84-86, 89-92,94,96-102, 104,105,107,109, 110,124
KIAA0092	CEP57	(A)8	11q21	7 %	3/43	7 %	55
KIAA0295	ZNF609	(A)8	15q22.31	8 %	3/39	8 %	55
KIAA0335	ZNF518	(A)9	10q24.1	7 %	3/43	7 %	55
KIAA0336	GCC2	(A)8	2q12.3	8 %	3/43	7 %	55,56
KIAA0355	KIAA0355	(A)9	19q13.11	4 %	1/24	4 %	111
KIAA0530	ZNF292	(A)9	6q15	7 %	3/44	7 %	55

TABLE I, continued

Mutation frequencies of microsatellite-containing genes

Gene	HGNC	Repeat	Chr. Location	Mut % (tot)	Mut. S/C	Mut %(S/C)	Refs.
KIAA0595	PPRC1	(C)8	10q24.32	7 %	3/43	7 %	[55]
KIAA0754		(A)8	1p34.3	10 %	4/41	10 %	[55]
KIAA0844	ZNF365	(A)8	10q21.2	9 %	4/44	9 %	[55]
KIAA0905	SEC31A	(A)9	4q21.22	17 %	6/43	14 %	[55,56]
KIAA0943	ATG4B	(T)9	2q37.3	11 %	4/44	9 %	[55,56]
KIAA0977	COBLL1	(T)9	2q24.3	20 %	10/42	24 %	[55,56]
KIAA1052	CEP164	(A)11	11q23.3	31 %	12/39	31 %	[44]
KIAA1268	PARP14	(A)10	3q21.1	23 %	9/39	23 %	[44]
KIAA1333	KIAA1333	(A)10	14q12	21 %	8/39	21 %	[44]
KIAA1470	RCC2	(A)10	1p36.13	46 %	18/39	46 %	[44]
KKIAMRE	CDKL2	(A)9	4q21.1	4 %	2/57	4 %	[49]
MAC30	TMEM97	(A)10	17q11.2	17 %	9/60	15 %	[44,53,56]
MARCKS	MARCKS	(A)11	6q22.2	74 %	42/58	72 %	[44,52,54,59]
MAZ	MAZ	(C)8	16p11.2	8 %	3/38	8 %	[105]
MBD4	MBD4	(A)10	3q21.3	24 %	76/384	20 %	[17,44,53,55,56,59, 62,67,99,125-128]
MCT4	SLC16A4	(T)9	1p12	15 %	4/36	11 %	[55,56]
MKI67	MKI67	(A)8	10q26.2	18 %			[56]
MLH3	MLH3	(A)9	14q24.3	8 %	4/27	15 %	[122,129]
MRE11	MRE11A	(T)11	11q21	75 %	55/64	86 %	[65,67,130-132]
MRP2	ABCC2	(A)8	10q24	8 %	3/38	8 %	[105]
MYO10	MYO10	(G)8	5p15.1	11 %	4/38	11 %	[105]
NBS1	NBN	(A)7	8q21.3	0 %	0/39	0 %	[64]
NDUFC2	NDUFC2	(T)9	11q14.1	31 %	12/43	28 %	[54-56]
NKTR	NKTR	(C)8	4q32.1	7 %	3/43	7 %	[55]
NSEP	YBX1	(C)8	1p34.2	0 %	0/82	0 %	[46,71,87]
OGT	OGT	(T)10	Xq13.1	22 %	26/116	22 %	[44,49,53]
P4HB	P4HB	(A)8	17q25.3	10 %	3/42	7 %	[55,56]
PA2G4	PA2G4	(A)8	12q13.2	18 %	9/43	21 %	[55,56]
PMS2	PMS2	(A)8	7p22.1	2 %	5/207	2 %	[66,71,87,111, 126]
POLA	POLA1	(A)8	Xp22.11	0 %	0/66	0 %	[71,111]
PRCC	PRCC	(C)8	1q21.1	12 %			[56]
PRKCI	PRKCI	(A)8	3q26.2	11 %			[56]
PRKWNK1	WNK1	(A)10	12p13.3	23 %	9/39	23 %	[44]
PRRG1	PRRG1	(C)8	Xp21.1	9 %	4/43	9 %	[55]
PTEN	PTEN	(A)6 * 2	10q23.31	17 %	26/138	19 %	[59,69,100,133-137]
PTHL3	PTHLH	(A)11	12p11.22	91 %	18/20	90 %	[53,54,59]
PTPN21	PTPN21	(A)8	14q31.3	13 %	5/43	12 %	[55,56]
RAB2L	RGL2	(G)8	6p21.3	12 %	5/43	12 %	[55]
RACK7	ZMYND8	(A)8	20q13.12	15 %	20/135	15 %	[55,66]
RBBP2	JARID1A	(A)8	12p11	17 %			[56]
RBBP8	RBBP8	(A)9	18q11.2	17 %	30/179	17 %	[49,61,115,138]
RECQL	RECQL	(A)9	12p12.1	8 %	19/213	9 %	[17,49,66,111]
RFC3	RFC3	(A)10	13q12.3-13	21 %	8/39	21 %	[44]
RGS12	RGS12	(C)8	4p16.3	29 %	11/38	29 %	[105]

TABLE I, continued
Mutation frequencies of microsatellite-containing genes

Gene	HGNC	Repeat	Chr. Location	Mut % (tot)	Mut. S/C	Mut %(S/C)	Refs.
RHAMM	HMMR	(A)9	5q34	16 %	9/57	16 %	[49]
RIP140	NRIP1	(A)9	21q11.2	9 %	3/42	7 %	[55,56]
RIS1	TMEM158	(GCN)14	3p21.31	44 %	7/16	44 %	[139]
RIZ	PRDM2	(A)8, (A)9	1p36.21	35 %	24/83	29 %	[56,59,111,140-142]
SEC63	SEC63	(A)10,(A)9	6q16-22	54 %	58/103	56 %	[44,53,55,56]
SEMG1	SEMG1	(T)9	20q13.12	51 %	74/146	51 %	[143]
SEX	PLXNA3	(G)8	Xq28	14 %	5/35	14 %	[55]
SHC1	SHC1	(G)8	1q22	0 %	0/6	0 %	[110]
SLC17A2	SLC17A2	(A)8	6p21.3	12 %			[56]
SLC23A1	SLC23A2	(C)9	20p13	45 %	9/20	45 %	[53]
SLC4A3	SLC4A3	(C)9	2q35	33 %	7/21	33 %	[53,54]
SPINK5	SPINK5	(A)10	5q32	31 %	12/39	31 %	[44]
SREBP2	SREBF2	(CAG)12	22q13.2	6 %	1/18	6 %	[119]
ß2m	B2M	(CT)4, 2*(A)6	15q21.1	29 %	5/17	29 %	[144,145]
STK11	STK11	(C)6	19p13.3	8 %	6/80	8 %	[146]
SYCP1	SYCP1	(A)10	1p13-12	17 %	11/60	18 %	[44,53,56]
TAF-1B	TAF1B	(A)11	2p25	78 %	45/58	78 %	[44,52]
TAN-1	NOTCH1	(ACC)6	9q34.3	11 %	2/18	11 %	[119]
TAP1	TAP1	(G)6	6p21.32	11 %	2/18	11 %	[145]
TAP2	TAP2	(C)6	6p21.32	12 %	1/17	6 %	[145]
TCF1	HNF1A	(C)8	12q24.3	32 %	12/38	32 %	[105]
TCF-4	TCF7L2	(A)9	10q25.2	36 %	126/307	41 %	[17,49,59,62,67,85, 105,147]
TCF6L1	TCF6L1	(A)10	7pter-cen	47 %	27/57	47 %	[44,52]
TEF4	TEAD2	(C)8	19q13.33	32 %	12/38	32 %	[54,105]
TFDP2	TFDP2	(A)8	3q23	3 %	0/57	0 %	[56,115]
TFE3	TFE3	(C)8	Xp11.22	24 %	9/38	24 %	[105]
TGF-ßRII	TGFBR2	(A)10	3p24.1	71 %	759/951	80 %	[17,44,49,53,55,56,59,62,67,70,73-75, 77,80,83-86,88-94, 96,98-105,108, 109,136,148-154]
TLOC1	TLOC1	(A)9	3q26.2	7 %	4/57	7 %	[49]
TPRDI	TTC3	(A)8	21q22-13	11 %	4/44	9 %	[55,56]
TSC1	TSC1	(GCA)6	9q34.13	0 %	0/6	0 %	[110]
TTK	TTK	(A)9	6q14.1	28 %			[56]
USP-1	USP1	(A)8	1p31.3	17 %			[56]
UVRAG	UVRAG	(A)10	11q13.5	35 %	20/57	35 %	[44,52]
VRK2	VRK2	(A)8	2p16.1	11 %			[56]
WBP1	WBP1	(C)9	2p13.1	9 %	3/43	7 %	[55,56]
WISP3	WISP3	(A)9	6q21	22 %	11/36	31 %	[59,62]
WRN	WRN	(A)8	8p12	0 %	0/6	0 %	[110]
XPOT	XPOT	(T)9	12q14.2	14 %	6/43	14 %	[55]
ZFP103	RNF103	(A)8	2p11.2	20 %			[56]

*The mutation frequency for CBP is pooled with that of EP300; the dimorphic trinucleotide repeat of E2F4 is (CAG)12-13.

Microsatellites that have short repeat units of 1-3 bp and those with the highest repeat number (e.g., mononucleotide runs of over 7 bp) are most prone to replication slippage,[155] and of the target genes given; those containing coding mononucleotide repeats (cMNRs) of (N)≥8 outnumber the others six to one, with (A)8-10 being the most common. (A/T) cMNRs of over 16bp are exceptionally rare in the genome; (C) and (G) cMNRs do not exceed 16bp and 13bp, respectively. Certain genes contain more than one cMNR, many contain multiple (N)6s, but there are indications that only one repeat is subject to most of the mutaional events.[52] Di- and trinucleotide repeats are infrequently represented among the target genes. This may be attributed to a lower inherent propensity for replication slippage, and for the trinucleotide tracts the assumption that insertion or deletion of a full unit will make only a potentially marginal difference in the protein product as it will be in-frame. However, the addition or deletion of an amino acid can have a profound effect on a protein, depending on size, charge, location, etc., and not all the indels of trinucleotide coding repeats are of one repeat unit. *E2F4*, for example, the gene containing a (CAG) repeat that is polymorphic in normal tissue,[94,117] appears to enhance proliferation when it contains inserted or deleted codons,[156] and the imperfect triplet repeat in *RIS1* is subject to frameshift mutations which interrupt its product's polyalanine domain.[139]

VI. TARGET GENES: DEFINITIONS AND EFFECTS

In theory, the random nature of mutation would mean that each microsatellite has the same likelihood of being hit, provided there are no sequence-dependent structural features which affect the basal replication error rate. So, given MMR deficiency, there will be a background rate of microsatellite mutation, whereas those short tandem repeat-containing genes which are truly involved in tumorigenesis should be found to have a significantly higher mutation rate.[11] The background level has been estimated to be in the area of 10-15%.[49,157] It is also generally assumed that the carcinogenetic potential of MSI target genes depends on said genes having the relevant oligonucleotide repeats in a coding sequence, though microsatellites in introns near exon boundaries and in UTRs have also been examined for elevated mutation frequencies due to their putative roles in splicing and transcriptional regulation,[158] respectively.

Several schemes have been put forward for the definition and classification of target genes of mismatch repair defective cancers, initially by the National Cancer Institute meeting of 1997[7] and subsequently by Duval and Hamelin, and Woerner et al.[52,159] According to the NCI criteria a true target gene must (1) have a high frequency of inactivation, (2) be subject to biallelic inactivation, (3) take part in a defined growth suppressor pathway, (4) the same growth suppressor pathway as above must exhibit inactivation in MSS tumors, and finally, (5) the

gene must be validated by functional suppressor studies in in vivo or in vitro models. These criteria have been criticized as excessively narrow:[160] (1) Inactivation events are not the only functionally important types of mutation, for example, *AXIN2* and *TCF4* are acknowledged target genes but the mutations to which they are subject are not expected to be inactivating. (2) Biallelic inactivation need not be a requirement in the case of haploinsufficiency (see below), c.f. the dearth of biallelic *TAF1B* inactivation.[44] To be truly useful, 3 and 4 would have to entail complete knowledge both of all possible growth suppressor pathways and of *every* gene or pathway involved mismatch repair proficient cancers, and 4 includes the assumption that the molecular pathway to MSS and MSI tumors are essentially equivalent, which is not necessarily the case. Nor do all acknowledged target genes participate in growth suppressor pathways, most notably *MSH3/6*.[159] Regarding criterion 5, such studies are lacking for most target genes, and furthermore, it is unsuitable for several types of potential target genes. The transformed phenotype, for example, will not be reversed in the event of reintroduction of a wild-type mutator target gene to a system,[160] and a gene with no known *functional* significance in tumor progression may yet be of prognostic clinical significance. Among the genes which have been subject to functional studies are *AXIN2*, *BAX, E2F4*, *RIZ*, *TCF4*, and *TGFβRII*.[45,68,111,156,161,162]

Duval and Hamelin proposed a fourfold, functional classification of affected genes into survivor genes, hibernator genes, cooperator genes and transformator genes.[159]

Survivor genes encode vital products and whose inactivation should exert a negative selection pressure. Hibernator genes are nonvital and downregulated, and should have a mutation rate in the background range. Cooperator genes designate sets of genes with the same terminal effect, e.g., promotion of apoptosis, which have a synergistic effect without any one gene requiring a high mutation frequency. Transformator genes are those which, upon mutation, independently confer a selection advantage to the cells concerned, and therefore should have the highest mutation frequency. These categories are thought to be mutated in a preferential order, with the transformator *TGFβRII* being among the earliest.[49,163]

The statistical regression model presented by Woerner *et al*.[52] takes into consideration the fact that longer repetitive tracts are more mutable (see below), i.e. the background rate for them is higher, and a gene with a mutation frequency above the 95% prediction interval for any given repeat length is considered a real target gene. *TGFβRII, BAX, TCF4, MSH3, ACVR2, PTHL3, HT001,* and *SLC23A1* are, by this method, considered genuine positive targets for MSI colorectal cancer, while the authors acknowledge the inapplicability of the model as regards target pathways, c.f. cooperator genes above.

Due to the difficulties of implementing clear-cut qualitative criteria—functional aspects of single genetic products and their interactions are often insufficiently elucidated, likewise signaling pathways and cascades—it has been most common to use the unmanipulated mutation frequency as the primary or

even sole criterion for target gene detection, and to treat any involvement of a frequently mutated gene in, e.g., apoptosis or cell cycle control as a bonus. Another potentially complicating factor is that, however detrimental frameshifts usually are, mutations in the repetitive tracts of target genes do not invariably cause complete inactivation. This appears to be the case for *AXIN2*, where the mutated product is more stable than wild type and may have a dominant negative effect,[68] and also for the mutated isoform of TCF4.[162] Both mutated gene products encourage inappropriate WNT signaling activation, which is cancer-promoting.[164]

The very fact that so many different mutational constellations exist suggests that the are few, if any, truly key genes for carcinogenesis among the target genes, *TGFβRII* being the only one to have been accorded such status.[44,55,159] Rather, the cumulative effect of many different and interchangeable mutations may drive tumorigenesis, with very few of the total being decisive in themselves.[159] Nor is it likely that all the relevant microsatellite-containing genes have as yet been tested for mutations in MSI tumors, or even necessarily characterized in any DNA sequence database. Several studies have used genome-wide sequence database searches for genes containing cMNRs as a basis for potential target gene selection.[44,53,55,105,119] Such a search* currently yields well over 1000 protein-coding genes containing the most promising (N)≥8 repeats, the figure rising more than tenfold when the range is expanded to include (N)6-7. Even allowing for unpublished negative data, it may be seen from the number of genes in Table I that fundamental target genes may still lurk in the unplumbed depths of the human genome.

VII. PERSPECTIVES

In addition to the mutational frequency and association studies of small numbers of target genes, future work will hopefully determine the prognostic values of distinct combinations of target genes—the prognostic value of any target gene mutations is at present unclear. None of the studies examining the matter consistently corroborate each other, despite being almost solely concerned with *BAX* and *TGFβRII* mutations and their inferred effect on patient survival ranges from a poor prognosis through prognostically irrelevant to improved survival.[108,154,161,166,167]

* **Scanning for monorepeats in human genes.** The 41,030 coding sequences in the transcripts of 20,484 human protein-coding genes were downloaded using the BioMart service at www.biomart.org [165] on March 13, 2007. A Perl script was written to scan the sequences for repeats. For each gene, only the longest coding sequence was considered. The script identified all mononucleotide repeats of length six and over, and also produced summary information about the repeats in each gene (longest repeat, number of repeats, sum of length of repeats).

Another topic that deserves attention is the functional consequence of mutation of mutated target gene expression. Both on the transcriptomic and proteomic levels, a large-scale approach to this area would be useful in understanding the role of MSI in oncogenesis.

Considering the comparative wealth of untested putative MMR target genes among cMNR-containing sequences alone, not to mention those harboring di- and trinucleotide or repeats in UTRs and introns, it seems clear that a high-throughput approach to analysis is desirable. Ideally, if somewhat unrealistically, all potential targets—both those that have been assessed in MSI tumors and those that have not—should be evaluated for a large consecutive clinical series.

In short, after 12 years of research into targeted genes of mismatch repair deficient colorectal cancers, we believe that those identified scarcely constitute even the tip of the iceberg.

REFERENCES

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, Thun MJ. Cancer statistics, 2006. CA Cancer J Clin. 2006;56:106–30.
2. Radtke F, Clevers H. Self-renewal and cancer of the gut: two sides of a coin. Science. 2005;307(5717):1904–9.
3. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61:759–67.
4. Grady WM. Genomic instability and colon cancer. Cancer Metastasis Rev. 2004;23:11–27.
5. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976; 194(4260):23–8.
6. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature. 1998;396(6712):643–9.
7. 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;58:5248–57.
8. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. Nature. 1997;386(6625):623–7.
9. Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C. The significance of unstable chromosomes in colorectal cancer. Nat Rev Cancer. 2003;3:695–701.

10. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science. 1993;260(5109):816–9.
11. Duval A, Hamelin R. Genetic instability in human mismatch repair deficient cancers. Ann Genet. 2002;45:71–5.
12. 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(5109):812–6.
13. Lothe RA, Peltomaki P, Meling GI, Aaltonen LA, Nystrom-Lahti M, Pylkkanen L, Heimdal K, Andersen TI, Moller P, Rognum TO, . Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. Cancer Res. 1993;53:5849–52.
14. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature. 1993;363(6429):558–61.
15. Bubb VJ, Curtis LJ, Cunningham C, Dunlop MG, Carothers AD, Morris RG, White S, Bird CC, Wyllie AH. Microsatellite instability and the role of hMSH2 in sporadic colorectalcancer. Oncogene. 1996;20;12:2641–9.
16. Kim H, Jen J, Vogelstein B, Hamilton SR. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. Am J Pathol. 1994;145:148–56.
17. Chung KY, Kim NG, Li LS, Kim H, Kim H, Nam CM, Kim H, Shin DH. Clinicopathologic characteristics related to the high variability of coding mononucleotide repeat sequences in tumors with high-microsatellite instability. Oncol Rep. 2003;10:439–44.
18. Glebov OK, Rodriguez LM, Nakahara K, Jenkins J, Cliatt J, Humbyrd CJ, DeNobile J, Soballe P, Simon R, Wright G, Lynch P, Patterson S, Lynch H, Gallinger S, Buchbinder A, Gordon G, Hawk E, Kirsch IR. Distinguishing right from left colon by the pattern of gene expression. Cancer Epidemiol Biomarkers Prev. 2003;12:755–62.
19. Bufill JA. Colorectal cancer: evidence for distinct genetic categories based on proximal or distal tumor location. Ann Intern Med. 1990;113:779–88.
20. Pocard M, Salmon RJ, Muleris M, Remvikos Y, Bara J, Dutrillaux B, Poupon MF. Deux colons—deux cancers? Adenocarcinomes coliques proximal ou distal: arguments en faveur d'une cancerogenese distincte [Two colons--two cancers? Proximal or distal adenocarcinoma: arguments for a different carcinogenesis]. Bull Cancer. 1995;82:10–21.
21. 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;96:8681–6.

22. Bodmer WF. Cancer genetics: colorectal cancer as a model. J Hum Genet. 2006;51:391–6.

23. Peltomaki P, Aaltonen LA, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Green JS, Jass JR, Weber JL, Leach FS, . Genetic mapping of a locus predisposing to human colorectal cancer. Science. 1993;260(5109):810–2.

24. Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell. 1993;75:1027–38.

25. Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Peltomaki P, Sistonen P, Aaltonen LA, Nystrom-Lahti M, . Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell. 1993;75:1215–25.

26. Watson P, Riley B. The tumor spectrum in the Lynch syndrome. Fam Cancer. 2005;4:245–8.

27. 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;96:261–8.

28. de la Chapelle A. Testing tumors for microsatellite instability. Eur J Hum Genet. 1999;7:407–8.

29. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol. 2005;20;23:609–18.

30. Sinicrope FA, Rego RL, Halling KC, Foster N, Sargent DJ, La PB, French AJ, Laurie JA, Goldberg RM, Thibodeau SN, Witzig TE. Prognostic impact of microsatellite instability and DNA ploidy in human colon carcinoma patients. Gastroenterology. 2006;131:729–37.

31. Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. Frameshift mutations and the genetic code. Cold Spring Harb Symp Quant Biol. 1966;31:77–84.

32. Kunkel TA. Misalignment-mediated DNA synthesis errors. Biochemistry. 1990;29:8003–11.

33. Perucho M. Cancer of the microsatellite mutator phenotype. Biol Chem. 1996; 377:675–84.

34. 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;57:808–11.

35. Kuismanen SA, Holmberg MT, Salovaara R, de la CA, Peltomaki P. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. Am J Pathol. 2000;156:1773–9.

36. Thibodeau SN, French AJ, Cunningham JM, Tester D, Burgart LJ, Roche PC, McDonnell SK, Schaid DJ, Vockley CW, Michels VV, Farr GH, Jr., O'Connell MJ. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. Cancer Res. 1998;58:1713–8.

37. 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;58:3455–60.

38. 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;95:6870–5.

39. Borresen AL, Lothe RA, Meling GI, Lystad S, Morrison P, Lipford J, Kane MF, Rognum TO, Kolodner RD. Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. Hum Mol Genet. 1995;4:2065–72.

40. Liu B, Nicolaides NC, Markowitz S, Willson JK, Parsons RE, Jen J, Papadopolous N, Peltomaki P, de la CA, Hamilton SR, . Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. Nat Genet. 1995;9:48–55.

41. Jascur T, Boland CR. Structure and function of the components of the human DNA mismatch repair system. Int J Cancer. 2006;119:2030–5.

42. Marti TM, Kunz C, Fleck O. DNA mismatch repair and mutation avoidance pathways. J Cell Physiol. 2002;191:28–41.

43. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.

44. Kim NG, Rhee H, Li LS, Kim H, Lee JS, Kim JH, Kim NK, Kim H. Identification of MARCKS, FLJ11383 and TAF1B as putative novel target genes in colorectal carcinomas with microsatellite instability. Oncogene. 2002;21:5081–7.

45. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, Fan RS, Zborowska E, Kinzler KW, Vogelstein B, . Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science. 1995; 268(5215):1336–8.

46. Rampino N, Yamamoto H, Ionov Y, Li Y, Sawai H, Reed JC, Perucho M. Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. Science. 1997;275(5302):967–9.

47. Souza RF, Appel R, Yin J, Wang S, Smolinski KN, Abraham JM, Zou TT, Shi YQ, Lei J, Cottrell J, Cymes K, Biden K, Simms L, Leggett B, Lynch PM, Frazier M, Powell SM, Harpaz N, Sugimura H, Young J, Meltzer SJ. Microsatel-

lite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumours. Nat Genet. 1996;14:255–7.

48. Malkhosyan S, Rampino N, Yamamoto H, Perucho M. Frameshift mutator mutations. Nature. 1996;382(6591):499–500.
49. Duval A, Rolland S, Compoint A, Tubacher E, Iacopetta B, Thomas G, Hamelin R. Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers. Hum Mol Genet. 2001;10:513–8.
50. Perucho M. Microsatellite instability: the mutator that mutates the other mutator. Nat Med. 1996;2:630–-1.
51. Vo AT, Zhu F, Wu X, Yuan F, Gao Y, Gu L, Li GM, Lee TH, Her C. hMRE11 deficiency leads to microsatellite instability and defective DNA mismatch repair. EMBO Rep. 2005;6:438–44.
52. Woerner SM, Benner A, Sutter C, Schiller M, Yuan YP, Keller G, Bork P, Doeberitz MK, Gebert JF. Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. Oncogene. 2003;22:2226–35.
53. Woerner SM, Gebert J, Yuan YP, Sutter C, Ridder R, Bork P, von Knebel DM. Systematic identification of genes with coding microsatellites mutated in DNA mismatch repair-deficient cancer cells. Int J Cancer. 2001;93:12–9.
54. Woerner SM, Kloor M, Mueller A, Rueschoff J, Friedrichs N, Buettner R, Buzello M, Kienle P, Knaebel HP, Kunstmann E, Pagenstecher C, Schackert HK, Moslein G, Vogelsang H, von Knebel DM, Gebert JF. Microsatellite instability of selective target genes in HNPCC-associated colon adenomas. Oncogene. 2005;24:2525–35.
55. Mori Y, Yin J, Rashid A, Leggett BA, Young J, Simms L, Kuehl PM, Langenberg P, Meltzer SJ, Stine OC. Instabilotyping: comprehensive identification of frameshift mutations caused by coding region microsatellite instability. Cancer Res. 2001;61:6046–9.
56. Mori Y, Sato F, Selaru FM, Olaru A, Perry K, Kimos MC, Tamura G, Matsubara N, Wang S, Xu Y, Yin J, Zou TT, Leggett B, Young J, Nukiwa T, Stine OC, Abraham JM, Shibata D, Meltzer SJ. Instabilotyping reveals unique mutational spectra in microsatellite-unstable gastric cancers. Cancer Res. 2002;62:3641–5.
57. Jung B, Doctolero RT, Tajima A, Nguyen AK, Keku T, Sandler RS, Carethers JM. Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. Gastroenterology. 2004;126:654–9.
58. Schulmann K, Mori Y, Croog V, Yin J, Olaru A, Sterian A, Sato F, Wang S, Xu Y, Deacu E, Berki AT, Hamilton JP, Kan T, Abraham JM, Schmiegel W, Harpaz N, Meltzer SJ. Molecular phenotype of inflammatory bowel disease-

associated neoplasms with microsatellite instability. Gastroenterology. 2005;129:74–85.

59. Yamaguchi T, Iijima T, Mori T, Takahashi K, Matsumoto H, Miyamoto H, Hishima T, Miyaki M. Accumulation profile of frameshift mutations during development and progression of colorectal cancer from patients with hereditary nonpolyposis colorectal cancer. Dis Colon Rectum. 2006;49:399–406.
60. Yamamoto H, Gil J, Schwartz S Jr, Perucho M. Frameshift mutations in Fas, Apaf-1, and Bcl-10 in gastro-intestinal cancer of the microsatellite mutator phenotype. Cell Death Differ. 2000;7:238–9.
61. Ikenoue T, Togo G, Nagai K, Ijichi H, Kato J, Yamaji Y, Okamoto M, Kato N, Kawabe T, Tanaka A, Matsumura M, Shiratori Y, Omata M. Frameshift mutations at mononucleotide repeats in RAD50 recombinational DNA repair gene in colorectal cancers with microsatellite instability. Jpn J Cancer Res. 2001;92:587–91.
62. Thorstensen L, Holm R, Lothe RA, Trope C, Carvalho B, Sobrinho-Simoes M, Seruca R. WNT-inducible signaling pathway protein 3, WISP-3, is mutated in microsatellite unstable gastrointestinal carcinomas but not in endometrial carcinomas. Gastroenterology. 2003;124:270–1.
63. Ejima Y, Yang L, Sasaki MS. Aberrant splicing of the ATM gene associated with shortening of the intronic mononucleotide tract in human colon tumor cell lines: a novel mutation target of microsatellite instability. Int J Cancer. 2000;86:262–8.
64. Kim NG, Choi YR, Baek MJ, Kim YH, Kang H, Kim NK, Min JS, Kim H. Frameshift mutations at coding mononucleotide repeats of the hRAD50 gene in gastrointestinal carcinomas with microsatellite instability. Cancer Res. 2001;61:36–8.
65. Ham MF, Takakuwa T, Luo WJ, Liu A, Horii A, Aozasa K. Impairment of double-strand breaks repair and aberrant splicing of ATM and MRE11 in leukemia-lymphoma cell lines with microsatellite instability. Cancer Sci. 2006;97:226–34.
66. Park J, Betel D, Gryfe R, Michalickova K, Di NN, Gallinger S, Hogue CW, Redston M. Mutation profiling of mismatch repair-deficient colorectal cncers using an in silico genome scan to identify coding microsatellites. Cancer Res. 2002;62:1284–8.
67. Mongiat-Artus P, Miquel C, Van der AM, Buhard O, Hamelin R, Soliman H, Bangma C, Janin A, Teillac P, van der KT, Praz F. Microsatellite instability and mutation analysis of candidate genes in urothelial cell carcinomas of upper urinary tract. Oncogene. 2006;25:2113–8.
68. Liu W, Dong X, Mai M, Seelan RS, Taniguchi K, Krishnadath KK, Halling KC, Cunningham JM, Boardman LA, Qian C, Christensen E, Schmidt SS,

Roche PC, Smith DI, Thibodeau SN. Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. Nat Genet. 2000;26:146–7.

69. Thorstensen L, Lind GE, Lovig T, Diep CB, Meling GI, Rognum TO, Lothe RA. Genetic and epigenetic changes of components affecting the WNT pathway in colorectal carcinomas stratified by microsatellite instability. Neoplasia. 2005;7:99–108.
70. Chung YJ, Park SW, Song JM, Lee KY, Seo EJ, Choi SW, Rhyu MG. Evidence of genetic progression in human gastric carcinomas with microsatellite instability. Oncogene. 1997;15:1719–26.
71. Yamamoto H, Sawai H, Perucho M. Frameshift somatic mutations in gastrointestinal cancer of the microsatellite mutator phenotype. Cancer Res. 1997;57:4420–6.
72. Catasus L, Matias-Guiu X, Machin P, Munoz J, Prat J. BAX somatic frameshift mutations in endometrioid adenocarcinomas of the endometrium: evidence for a tumor progression role in endometrial carcinomas with microsatellite instability. Lab Invest. 1998;78:1439–44.
73. Fujiwara T, Stolker JM, Watanabe T, Rashid A, Longo P, Eshleman JR, Booker S, Lynch HT, Jass JR, Green JS, Kim H, Jen J, Vogelstein B, Hamilton SR. Accumulated clonal genetic alterations in familial and sporadic colorectal carcinomas with widespread instability in microsatellite sequences. Am J Pathol. 1998;153:1063–78.
74. Oliveira C, Seruca R, Seixas M, Sobrinho-Simoes M. The clinicopathological features of gastric carcinomas with microsatellite instability may be mediated by mutations of different "target genes": a study of the TGFbeta RII, IGFII R, and BAX genes. Am J Pathol. 1998;153:1211–9.
75. Ottini L, Falchetti M, D'Amico C, Amorosi A, Saieva C, Masala G, Frati L, Cama A, Palli D, Mariani-Costantini R. Mutations at coding mononucleotide repeats in gastric cancer with the microsatellite mutator phenotype. Oncogene. 1998;16:2767–72.
76. Ouyang H, Furukawa T, Abe T, Kato Y, Horii A. The BAX gene, the promoter of apoptosis, is mutated in genetically unstable cancers of the colorectum, stomach, and endometrium. Clin Cancer Res. 1998;4:1071–4.
77. Percesepe A, Kristo P, Aaltonen LA, Ponz de LM, de la CA, Peltomaki P. Mismatch repair genes and mononucleotide tracts as mutation targets in colorectal tumors with different degrees of microsatellite instability. Oncogene. 1998;17:157–63.
78. Sakao Y, Noro M, Sekine S, Nozue M, Hirohashi S, Itoh T, Noguchi M. Microsatellite instability and frameshift mutations in the Bax gene in hereditary nonpolyposis colorectal carcinoma. Jpn J Cancer Res. 1998;89:1020–7.

79. Simms LA, Radford-Smith G, Biden KG, Buttenshaw R, Cummings M, Jass JR, Young J, Meltzer SJ, Leggett BA. Reciprocal relationship between the tumor suppressors p53 and BAX in primary colorectal cancers. Oncogene. 1998;17:2003–8.

80. Wu MS, Lee CW, Shun CT, Wang HP, Lee WJ, Sheu JC, Lin JT. Clinicopathological significance of altered loci of replication error and microsatellite instability-associated mutations in gastric cancer. Cancer Res. 1998;58: 1494–7.

81. Yagi OK, Akiyama Y, Nomizu T, Iwama T, Endo M, Yuasa Y. Proapoptotic gene BAX is frequently mutated in hereditary nonpolyposis colorectal cancers but not in adenomas. Gastroenterology. 1998;114:268–74.

82. Yamamoto H, Sawai H, Weber TK, Rodriguez-Bigas MA, Perucho M. Somatic frameshift mutations in DNA mismatch repair and proapoptosis genes in hereditary nonpolyposis colorectal cancer. Cancer Res. 1998;58:997–1003.

83. Abdel-Rahman WM, Georgiades IB, Curtis LJ, Arends MJ, Wyllie AH. Role of BAX mutations in mismatch repair-deficient colorectal carcinogenesis. Oncogene. 1999;18:2139–42.

84. Bertoni F, Codegoni AM, Furlan D, Tibiletti MG, Capella C, Broggini M. CHK1 frameshift mutations in genetically unstable colorectal and endometrial cancers. Genes Chromosomes Cancer. 1999;26:176–80.

85. Duval A, Iacopetta B, Ranzani GN, Lothe RA, Thomas G, Hamelin R. Variable mutation frequencies in coding repeats of TCF-4 and other target genes in colon, gastric and endometrial carcinoma showing microsatellite instability. Oncogene. 1999;18:6806–9.

86. Iacopetta BJ, Soong R, House AK, Hamelin R. Gastric carcinomas with microsatellite instability: clinical features and mutations to the TGF-beta type II receptor, IGFII receptor, and BAX genes. J Pathol. 1999;187:428–32.

87. Schwartz S Jr, Yamamoto H, Navarro M, Maestro M, Reventos J, Perucho M. Frameshift mutations at mononucleotide repeats in caspase-5 and other target genes in endometrial and gastrointestinal cancer of the microsatellite mutator phenotype. Cancer Res. 1999;59:2995–3002.

88. Abe Y, Masuda H. Genetic alterations of sporadic colorectal cancer with microsatellite instability, especially characteristics of primary multiple colorectal cancers. J Surg Oncol. 2000;74:249–56.

89. Barnetson R, Jass J, Tse R, Eckstein R, Robinson B, Schnitzler M. Mutations associated with microsatellite unstable colorectal carcinomas exhibit widespread intratumoral heterogeneity. Genes Chromosomes Cancer. 2000; 29:130–6.

90. Calin GA, Gafa R, Tibiletti MG, Herlea V, Becheanu G, Cavazzini L, Barbanti-Brodano G, Nenci I, Negrini M, Lanza G. Genetic progression in microsatellite

instability high (MSI-H) colon cancers correlates with clinico-pathological parameters: A study of the TGRbetaRII, BAX, hMSH3, hMSH6, IGFIIR and BLM genes. Int J Cancer. 2000;89:230–5.

91. Catasus L, Matias-Guiu X, Machin P, Zannoni GF, Scambia G, edetti-Panici P, Prat J. Frameshift mutations at coding mononucleotide repeat microsatellites in endometrial carcinoma with microsatellite instability. Cancer. 2000;88: 2290–7.
92. Johannsdottir JT, Jonasson JG, Bergthorsson JT, Amundadottir LT, Magnusson J, Egilsson V, Ingvarsson S. The effect of mismatch repair deficiency on tumourigenesis; microsatellite instability affecting genes containing short repeated sequences. Int J Oncol. 2000;16:133–9.
93. Percesepe A, Pedroni M, Sala E, Menigatti M, Borghi F, Losi L, Viel A, Genuardi M, Benatti P, Roncucci L, Peltomaki P, Ponz de LM. Genomic instability and target gene mutations in colon cancers with different degrees of allelic shifts. Genes Chromosomes Cancer. 2000;27:424–9.
94. Semba S, Ouyang H, Han SY, Kato Y, Horii A. Analysis of the candidate target genes for mutation in microsatellite instability-positive cancers of the colorectum, stomach, and endometrium. Int J Oncol. 2000;16:731–7.
95. Simms LA, Young J, Wicking C, Meltzer SJ, Jass JR, Leggett BA. The apoptotic regulatory gene, BCL10, is mutated in sporadic mismatch repair deficient colorectal cancers. Cell Death Differ. 2000;7:236–7.
96. Calin G, Ranzani GN, Amadori D, Herlea V, Matei I, Barbanti-Brodano G, Negrini M. Somatic frameshift mutations in the Bloom syndrome BLM gene are frequent in sporadic gastric carcinomas with microsatellite mutator phenotype. BMC Genet. 2001;2:14. Epub. 2001 Aug 14.14.
97. Gras E, Catasus L, Arguelles R, Moreno-Bueno G, Palacios J, Gamallo C, Matias-Guiu X, Prat J. Microsatellite instability, MLH-1 promoter hypermethylation, and frameshift mutations at coding mononucleotide repeat microsatellites in ovarian tumors. Cancer. 2001;92:2829–36.
98. Lee HS, Lee BL, Kim SH, Woo DK, Kim HS, Kim WH. Microsatellite instability in synchronous gastric carcinomas. Int J Cancer. 2001;91:619–24.
99. Miyaki M, Iijima T, Shiba K, Aki T, Kita Y, Yasuno M, Mori T, Kuroki T, Iwama T. Alterations of repeated sequences in 5' upstream and coding regions in colorectal tumors from patients with hereditary nonpolyposis colorectal cancer and Turcot syndrome. Oncogene. 2001;20:5215–8.
100. Shin KH, Park YJ, Park JG. PTEN gene mutations in colorectal cancers displaying microsatellite instability. Cancer Lett. 2001;174:189–94.
101. Togo G, Shiratori Y, Okamoto M, Yamaji Y, Matsumura M, Sano T, Motojima T, Omata M. Relationship between grade of microsatellite instability and target

genes of mismatch repair pathways in sporadic colorectal carcinoma. Dig Dis Sci. 2001;46:1615–22.

102. Furlan D, Casati B, Cerutti R, Facco C, Terracciano L, Capella C, Chiaravalli AM. Genetic progression in sporadic endometrial and gastrointestinal cancers with high microsatellite instability. J Pathol. 2002;197:603–9.
103. Moriyama H, Sasamoto H, Kambara T, Matsubara N, Ikeda M, Baba S, Meltzer SJ, Lynch HT, Shimizu K, Tanaka N. E2F-4 mutation in hereditary nonpolyposis colorectal cancer. J Exp Clin Cancer Res. 2002;21:185–9.
104. Jeong SY, Shin KH, Shin JH, Ku JL, Shin YK, Park SY, Kim WH, Park JG. Microsatellite instability and mutations in DNA mismatch repair genes in sporadic colorectal cancers. Dis Colon Rectum. 2003;46:1069–77.
105. Potocnik U, Glavac D, Ravnik-Glavac M. Identification of novel genes with somatic frameshift mutations within coding mononucleotide repeats in colorectal tumors with high microsatellite instability. Genes Chromosomes Cancer. 2003;36:48–56.
106. Trojan J, Brieger A, Raedle J, Weber N, Kriener S, Kronenberger B, Caspary WF, Zeuzem S. BAX and caspase-5 frameshift mutations and spontaneous apoptosis in colorectal cancer with microsatellite instability. Int J Colorectal Dis. 2004;19:538–44.
107. Vassileva V, Millar A, Briollais L, Chapman W, Bapat B. Apoptotic and growth regulatory genes as mutational targets in mismatch repair deficient endometrioid adenocarcinomas of young patients. Oncol Rep. 2004;11:931–7.
108. Fernandez-Peralta AM, Nejda N, Oliart S, Medina V, Azcoita MM, Gonzalez-Aguilera JJ. Significance of mutations in TGFBR2 and BAX in neoplastic progression and patient outcome in sporadic colorectal tumors with high-frequency microsatellite instability. Cancer Genet Cytogenet. 2005;157:18–24.
109. Chen GT, Zhu ZG, Yin HR, Liu BY, Ji J, Zhang J, Lin YZ. [The relationship between frameshift mutations of transforming growth factor-beta type II receptor, insulin growth factor II receptor, bcl-2 associated X protein, E2F4 and microsatellite instability in gastric carcinoma]. Zhonghua Wai Ke Za Zhi. 2006;44:344–8.
110. Calin G, Herlea V, Barbanti-Brodano G, Negrini M. The coding region of the Bloom syndrome BLM gene and of the CBL proto-oncogene is mutated in genetically unstable sporadic gastrointestinal tumors. Cancer Res. 1998;58:3777–81.
111. Chadwick RB, Jiang GL, Bennington GA, Yuan B, Johnson CK, Stevens MW, Niemann TH, Peltomaki P, Huang S, de la CA. Candidate tumor suppressor RIZ is frequently involved in colorectal carcinogenesis. Proc Natl Acad Sci U S A. 2000;97:2662–7.

112. Ionov Y, Matsui S, Cowell JK. A role for p300/CREB binding protein genes in promoting cancer progression in colon cancer cell lines with microsatellite instability. Proc Natl Acad Sci U S A. 2004;101:1273–8.

113. Laghi L, Ranzani GN, Bianchi P, Mori A, Heinimann K, Orbetegli O, Spaudo MR, Luinetti O, Francisconi S, Roncalli M, Solcia E, Malesci A. Frameshift mutations of human gastrin receptor gene (hGARE) in gastrointestinal cancers with microsatellite instability. Lab Invest. 2002;82:265–71.

114. Wicking C, Simms LA, Evans T, Walsh M, Chawengsaksophak K, Beck F, Chenevix-Trench G, Young J, Jass J, Leggett B, Wainwright B. CDX2, a human homologue of Drosophila caudal, is mutated in both alleles in a replication error positive colorectal cancer. Oncogene. 1998;17:657–9.

115. El-Bchiri J, Buhard O, Penard-Lacronique V, Thomas G, Hamelin R, Duval A. Differential nonsense mediated decay of mutated mRNAs in mismatch repair deficient colorectal cancers. Hum Mol Genet. 2005;14:2435–42.

116. Yoshitaka T, Matsubara N, Ikeda M, Tanino M, Hanafusa H, Tanaka N, Shimizu K. Mutations of E2F-4 trinucleotide repeats in colorectal cancer with microsatellite instability. Biochem Biophys Res Commun. 1996;227:553–7.

117. Ikeda M, Orimo H, Moriyama H, Nakajima E, Matsubara N, Mibu R, Tanaka N, Shimada T, Kimura A, Shimizu K. Close correlation between mutations of E2F4 and hMSH3 genes in colorectal cancers with microsatellite instability. Cancer Res. 1998;58:594–8.

118. Alazzouzi H, Davalos V, Kokko A, Domingo E, Woerner SM, Wilson AJ, Konrad L, Laiho P, Espin E, Armengol M, Imai K, Yamamoto H, Mariadason JM, Gebert JF, Aaltonen LA, Schwartz S Jr, Arango D. Mechanisms of inactivation of the receptor tyrosine kinase EPHB2 in colorectal tumors. Cancer Res. 2005;65:10170–3.

119. Forgacs E, Wren JD, Kamibayashi C, Kondo M, Xu XL, Markowitz S, Tomlinson GE, Muller CY, Gazdar AF, Garner HR, Minna JD. Searching for microsatellite mutations in coding regions in lung, breast, ovarian and colorectal cancers. Oncogene. 2001;20:1005–9.

120. Swisher EM, Mutch DG, Herzog TJ, Rader JS, Kowalski LD, Elbendary A, Goodfellow PJ. Analysis of MSH3 in endometrial cancers with defective DNA mismatch repair. J Soc Gynecol Investig. 1998;5:210–6.

121. Orimo H, Nakajima E, Ikejima M, Emi M, Shimada T. Frameshift mutations and a length polymorphism in the hMSH3 gene and the spectrum of microsatellite instability in sporadic colon cancer. Jpn J Cancer Res. 1999;90:1310–5.

122. Akiyama Y, Nagasaki H, Nakajima T, Sakai H, Nomizu T, Yuasa Y. Infrequent frameshift mutations in the simple repeat sequences of hMLH3 in hereditary nonpolyposis colorectal cancers. Jpn J Clin Oncol. 2001;31:61–4.

123. Ohmiya N, Matsumoto S, Yamamoto H, Baranovskaya S, Malkhosyan SR, Perucho M. Germline and somatic mutations in hMSH6 and hMSH3 in gastrointestinal cancers of the microsatellite mutator phenotype. Gene. 2001;272: 301–13.

124. Ouyang H, Shiwaku HO, Hagiwara H, Miura K, Abe T, Kato Y, Ohtani H, Shiiba K, Souza RF, Meltzer SJ, Horii A. The insulin-like growth factor II receptor gene is mutated in genetically unstable cancers of the endometrium, stomach, and colorectum. Cancer Res. 1997;57:1851–4.

125. Bader S, Walker M, Hendrich B, Bird A, Bird C, Hooper M, Wyllie A. Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. Oncogene. 1999;18:8044–7.

126. Riccio A, Aaltonen LA, Godwin AK, Loukola A, Percesepe A, Salovaara R, Masciullo V, Genuardi M, Paravatou-Petsotas M, Bassi DE, Ruggeri BA, Klein-Szanto AJ, Testa JR, Neri G, Bellacosa A. The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. Nat Genet. 1999;23:266–8.

127. Yamada T, Koyama T, Ohwada S, Tago K, Sakamoto I, Yoshimura S, Hamada K, Takeyoshi I, Morishita Y. Frameshift mutations in the MBD4/MED1 gene in primary gastric cancer with high-frequency microsatellite instability. Cancer Lett. 2002;181:115–20.

128. Pinto M, Wu Y, Suriano G, Mensink RG, Duval A, Oliveira C, Carvalho B, Hamelin R, Seruca R, Hofstra RM. MBD4 mutations are rare in gastric carcinomas with microsatellite instability. Cancer Genet Cytogenet. 2003;145: 103–7.

129. Loukola A, Vilkki S, Singh J, Launonen V, Aaltonen LA. Germline and somatic mutation analysis of MLH3 in MSI-positive colorectal cancer. Am J Pathol. 2000;157:347–52.

130. Giannini G, Ristori E, Cerignoli F, Rinaldi C, Zani M, Viel A, Ottini L, Crescenzi M, Martinotti S, Bignami M, Frati L, Screpanti I, Gulino A. Human MRE11 is inactivated in mismatch repair-deficient cancers. EMBO Rep. 2002;3:248–54.

131. Ottini L, Falchetti M, Saieva C, De MM, Masala G, Zanna I, Paglierani M, Giannini G, Gulino A, Nesi G, Mariani CR, Palli D. MRE11 expression is impaired in gastric cancer with microsatellite instability. Carcinogenesis. 2004;25:2337–43.

132. Giannini G, Rinaldi C, Ristori E, Ambrosini MI, Cerignoli F, Viel A, Bidoli E, Berni S, D'Amati G, Scambia G, Frati L, Screpanti I, Gulino A. Mutations of an intronic repeat induce impaired MRE11 expression in primary human cancer with microsatellite instability. Oncogene. 2004;23:2640–7.

133. Guanti G, Resta N, Simone C, Cariola F, Demma I, Fiorente P, Gentile M. Involvement of PTEN mutations in the genetic pathways of colorectal cancerogenesis. Hum Mol Genet. 2000;9:283–7.

134. Bussaglia E, del RE, Matias-Guiu X, Prat J. PTEN mutations in endometrial carcinomas: a molecular and clinicopathologic analysis of 38 cases. Hum Pathol. 2000;31:312–7.

135. Cohn DE, Basil JB, Venegoni AR, Mutch DG, Rader JS, Herzog TJ, Gersell DJ, Goodfellow PJ. Absence of PTEN repeat tract mutation in endometrial cancers with microsatellite instability. Gynecol Oncol. 2000;79:101–6.

136. Dicuonzo G, Angeletti S, Garcia-Foncillas J, Brugarolas A, Okrouzhnov Y, Santini D, Tonini G, Lorino G, De CM, Baldi A. Colorectal carcinomas and PTEN/MMAC1 gene mutations. Clin Cancer Res. 2001;7:4049–53.

137. Zhou XP, Loukola A, Salovaara R, Nystrom-Lahti M, Peltomaki P, de la CA, Aaltonen LA, Eng C. PTEN mutational spectra, expression levels, and subcellular localization in microsatellite stable and unstable colorectal cancers. Am J Pathol. 2002;161:439–47.

138. Vilkki S, Launonen V, Karhu A, Sistonen P, Vastrik I, Aaltonen LA. Screening for microsatellite instability target genes in colorectal cancers. J Med Genet. 2002;39:785–9.

139. Iglesias D, Fernandez-Peralta AM, Nejda N, Daimiel L, Azcoita MM, Oliart S, Gonzalez-Aguilera JJ. RIS1, a gene with trinucleotide repeats, is a target in the mutator pathway of colorectal carcinogenesis. Cancer Genet Cytogenet. 2006;167:138–44.

140. Piao Z, Fang W, Malkhosyan S, Kim H, Horii A, Perucho M, Huang S. Frequent frameshift mutations of RIZ in sporadic gastrointestinal and endometrial carcinomas with microsatellite instability. Cancer Res. 2000;60:4701–4.

141. Sakurada K, Furukawa T, Kato Y, Kayama T, Huang S, Horii A. RIZ, the retinoblastoma protein interacting zinc finger gene, is mutated in genetically unstable cancers of the pancreas, stomach, and colorectum. Genes Chromosomes Cancer. 2001;30:207–11.

142. Pan KF, Lu YY, Liu WG, Zhang L, You WC. Detection of frameshift mutations of RIZ in gastric cancers with microsatellite instability. World J Gastroenterol. 2004;10:2719–22.

143. Hienonen T, Sammalkorpi H, Enholm S, Alhopuro P, Barber TD, Lehtonen R, Nupponen NN, Lehtonen H, Salovaara R, Mecklin JP, Jarvinen H, Koistinen R, Arango D, Launonen V, Vogelstein B, Karhu A, Aaltonen LA. Mutations in two short noncoding mononucleotide repeats in most microsatellite-unstable colorectal cancers. Cancer Res. 2005;65:4607–13.

144. Yamamoto H, Perez-Piteira J, Yoshida T, Terada M, Itoh F, Imai K, Perucho M. Gastric cancers of the microsatellite mutator phenotype display characteristic genetic and clinical features. Gastroenterology. 1999;116:1348–57.

145. Kloor M, Becker C, Benner A, Woerner SM, Gebert J, Ferrone S, von Knebel DM. Immunoselective pressure and human leukocyte antigen class I antigen machinery defects in microsatellite unstable colorectal cancers. Cancer Res. 2005;65:6418–24.

146. Nakagawa H, Koyama K, Nakamori S, Kameyama M, Imaoka S, Monden M, Nakamura Y. Frameshift mutation of the STK11 gene in a sporadic gastrointestinal cancer with microsatellite instability. Jpn J Cancer Res. 1999;90: 633–7.

147. Shimizu Y, Ikeda S, Fujimori M, Kodama S, Nakahara M, Okajima M, Asahara T. Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. Genes Chromosomes Cancer. 2002;33: 73–81.

148. Myeroff LL, Parsons R, Kim SJ, Hedrick L, Cho KR, Orth K, Mathis M, Kinzler KW, Lutterbaugh J, Park K. A transforming growth factor beta receptor type II gene mutation common in colon and gastric but rare in endometrial cancers with microsatellite instability. Cancer Res. 1995;55:5545–7.

149. Parsons R, Myeroff LL, Liu B, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. Cancer Res. 1995;55:5548–50.

150. Renault B, Calistri D, Buonsanti G, Nanni O, Amadori D, Ranzani GN. Microsatellite instability and mutations of p53 and TGF-beta RII genes in gastric cancer. Hum Genet. 1996;98:601–7.

151. Olschwang S, Hamelin R, Laurent-Puig P, Thuille B, De RY, Li YJ, Muzeau F, Girodet J, Salmon RJ, Thomas G. Alternative genetic pathways in colorectal carcinogenesis. Proc Natl Acad Sci U S A. 1997;94:12122–7.

152. Iacopetta BJ, Welch J, Soong R, House AK, Zhou XP, Hamelin R. Mutation of the transforming growth factor-beta type II receptor gene in right-sided colorectal cancer: relationship to clinicopathological features and genetic alterations. J Pathol. 1998;184:390–5.

153. Ohwada M, Suzuki M, Saga Y, Suzuki T, Ikeda M, Yamada M, Sato I. Mutational analysis of transforming growth factor beta receptor type II and DNA mismatch repair genes in sporadic endometrial carcinomas with microsatellite instability. Oncol Rep. 2000;7:789–92.

154. Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, Benson AB, III, Hamilton SR. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. N Engl J Med. 2001;19;344:1196–206.

155. Li YC, Korol AB, Fahima T, Beiles A, Nevo E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol. 2002;11:2453–65.

156. Takashima H, Matsumoto Y, Matsubara N, Shirakawa Y, Kawashima R, Tanino M, Ito S, Isozaki H, Ouchida M, Meltzer SJ, Shimizu K, Tanaka N. Effect of naturally occurring E2F-4 alterations on transcriptional activation and proliferation in transfected cells. Lab Invest. 2001;81:1565–73.

157. Duval A, Reperant M, Hamelin R. Comparative analysis of mutation frequency of coding and non coding short mononucleotide repeats in mismatch repair deficient colorectal cancers. Oncogene. 2002;21:8062–6.

158. Suraweera N, Iacopetta B, Duval A, Compoint A, Tubacher E, Hamelin R. Conservation of mononucleotide repeats within 3' and 5' untranslated regions and their instability in MSI-H colorectal cancer. Oncogene. 2001;20:7472–7.

159. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res. 2002;62:2447–54.

160. Perucho M. Correspondence re: C.R. Boland et al., 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., 58:5248–57, 1998. Cancer Res. 1999;59:249–56.

161. Ionov Y, Yamamoto H, Krajewski S, Reed JC, Perucho M. Mutational inactivation of the proapoptotic gene BAX confers selective advantage during tumor clonal evolution. Proc Natl Acad Sci U S A. 2000;97:10872–7.

162. Duval A, Rolland S, Tubacher E, Bui H, Thomas G, Hamelin R. The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. Cancer Res. 2000;60:3872–9.

163. Grady WM, Rajput A, Myeroff L, Liu DF, Kwon K, Willis J, Markowitz S. Mutation of the type II transforming growth factor-beta receptor is coincident with the transformation of human colon adenomas to malignant carcinomas. Cancer Res. 1998;58:3101–4.

164. Logan CY, Nusse R. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol. 2004;20:781–810.

165. Kasprzyk A, Keefe D, Smedley D, London D, Spooner W, Melsopp C, Hammond M, Rocca-Serra P, Cox T, Birney E. EnsMart: a generic system for fast and flexible access to biological data. Genome Res. 2004;14:160–9.

166. Samowitz WS, Curtin K, Neuhausen S, Schaffer D, Slattery ML. Prognostic implications of BAX and TGFBRII mutations in colon cancers with microsatellite instability. Genes Chromosomes Cancer. 2002;35:368–71.

167. Jung B, Smith EJ, Doctolero RT, Gervaz P, Alonso JC, Miyai K, Keku T, Sandler RS, Carethers JM. Influence of target gene mutations on survival, stage and histology in sporadic microsatellite unstable colon cancers. Int J Cancer. 2006;118:2509–13.