Genetic etiology of hereditary colorectal cancer: new mechanisms and advanced mutation detection techniques

Isabella Gazzoli

Genetic etiology of hereditary colorectal cancer: new mechanisms and advanced mutation detection techniques

Genetische etiologie van erfelijke dikkedarmkanker: nieuwe mechanismen en geavanceerde mutatie detectie technieken

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. W.H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op dinsdag 21 november 2006 des ochtends te 10.30 uur

door

Isabella Gazzoli

geboren op 5 maart 1970 te Milaan, Italië

Promotor: Prof. Dr. H.C. Clevers[1]

Co-promotor: Prof. Dr. R.D. Kolodner[2]

[1]Department of Physiological Chemistry
University Medical Center Utrecht
Utrecht, The Netherlands

[2]Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
UCSD School of Medicine
University of California, San Diego
La Jolla, USA

Nel mezzo del cammin di nostra vita
mi ritrovai per una selva oscura,
ché la diritta via era smarrita.

Ahi quanto a dir qual era è cosa dura
esta selva selvaggia e aspra e forte
che nel pensier rinova la paura!

Tant'è amara che poco è più morte;
ma per trattar del ben ch'i' vi trovai,
dirò de l'altre cose ch'i' v' ho scorte.

When I had journeyed half of our life's way,
I found myself within a shadowed forest,
for I had lost the path that does not stray.

Ah, it is hard to speak of what it was,
that savage forest, dense and difficult,
which even in recall renews my fear!

So bitter-death is hardly more severe;
but to retell the good discovered there,
I'll also tell the other things I saw.

(Comedia di Dante Alighieri, XIV Sec., Inferno, Canto I, 1-9, Dante online)

To my husband Arthur

Commissie: Prof. R. Fodde
Prof. R. H. Medema
Prof. G.J.A. Offerhaus
Prof. J. G.J. van de Winkel

Paranimfen: Dr. A.R. Houweling
Dr. J. de Rooij

Cover photo: Rheyngaerde by Jeroen van Vliet.

Printed by Optima Grafische Communicatie, Rotterdam
ISBN: 90-8559-231-3

Table of contents

Chapter 1

Introduction

1.1 Colorectal cancer

Transformation from normal cell to cancer cell is thought to be a multi-step process involving the accumulation of genetic alterations in oncogenes, tumor suppressor genes and stability genes. Mutations in oncogenes result in a gain of function by increasing the production of an unaltered protein or by creating a novel protein due to a chromosomal rearrangement. Mutations in oncogenes are somatic events caused by gene amplifications, chromosomal translocations or intragenic mutations affecting only a single allele, which is sufficient to provide a selective growth advantage to the cell. In contrast, mutations in tumor suppressor genes usually cause the inactivation of the genes resulting in the absence of the proteins. These genetic alterations include missense mutations (nucleotide substitutions resulting in amino acid change), nonsense mutations (nucleotide changes in the origin stop codon), partial or complete deletions of genes, insertions of nucleotides, epigenetic alterations (methylation of CpG sites) or modifications of transcription factor binding sites. According to the Knudson two-hit model of carcinogenesis, mutational events in both maternal and paternal alleles are required to cause a loss of function of tumor-suppressor genes [1, 2] or stability genes (also called caretakers). Stability genes maintain and regulate DNA fidelity. Inactivation of these genes results in genomic instability and increased rates of mutations in all genes. Neoplasms arise when tumor suppressor genes are inactivated and oncogenes are activated, or when an inefficient repair system results in increased rates of mutations in genes involved in cell growth or cell-cycle arrest.

Generally, tumors are classified as liquid or solid. The first type, which includes leukemia and lymphoma, occurs in mobile cells of the circulatory and lymphatic systems. Only one or two mutational events are sufficient to induce malignant transformation. These events typically involve oncogene activation by chromosome translocation. In contrast, three or more events are necessary to create a solid tumor, in epithelial or mesenchymal cells (immobile cells). Genetic alterations in oncogenes, tumor suppressor genes and stability genes can occur in germline or somatic cells, resulting in a hereditary predisposition or sporadic cancer, respectively. In the hereditary form, a mutation in one allele of the germline cell can be passed to the offspring after which a second mutation at the somatic level on the second allele can cause complete inactivation of the gene. In the sporadic form both alleles are inactivated at a somatic level during the life-time and cancer appears at a later age compared to the hereditary form.

1.1.1 Colorectal cancer is a good genetic model of tumorigenesis

In humans the large intestine is composed of the cecum, colon and rectum (Fig. 1). It is the last part of the digestive tract soon after the small intestine and measures about 1.5 meter. The colon is subdivided in the ascending side (right), the transverse, the descending side (left) and the sigmoid. After chemical digestion all nutrients are absorbed in the small intestine. When the chyme (unabsorbed nutrients) reaches the large intestine, only water, electrolytes and vitamins are removed, and the chyme becomes feces.

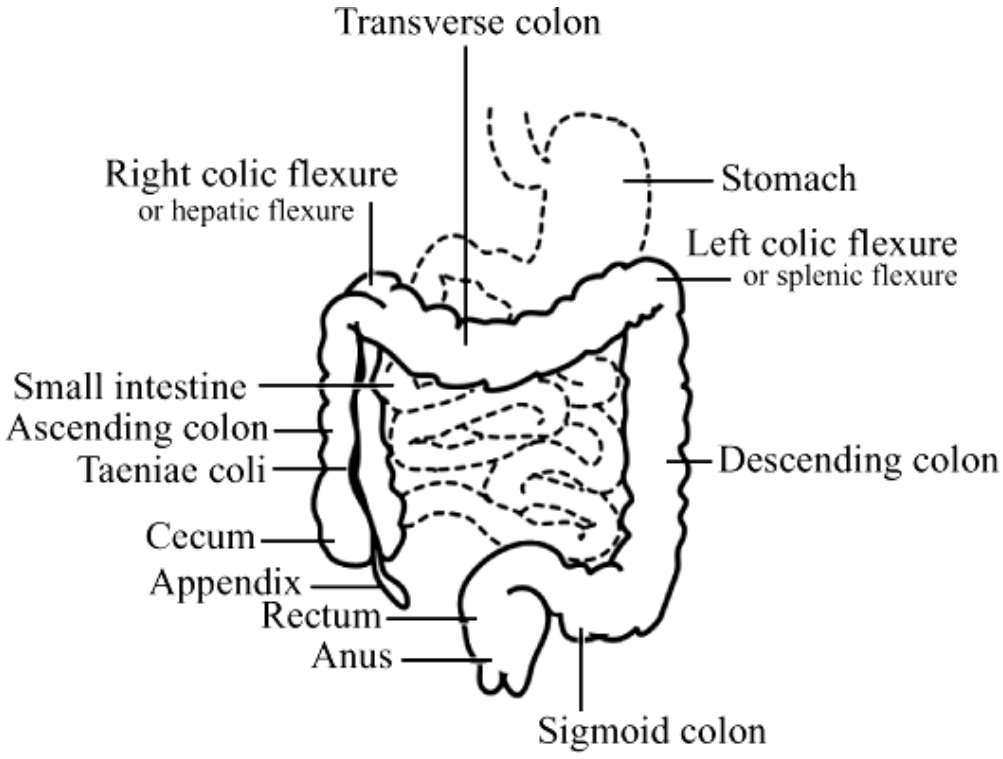

Figure 1. Partial representation of the human abdominal cavity containing stomach, small intestine, large intestine and anus. Large intestine is subdivided in cecum, colon (ascending, transverse, descending and sigmoid) and rectum.

The colon is a proliferating epithelium, which includes 10^7 compartments of cells, called crypts, each of which contains from 1000-4000 cells. The crypts in the lower part of the colon contain proliferative stem cells (making up two thirds of the total number of cells). The crypts in the upper third of the large intestine and the flat surface of the epithelium contain differentiated cells.

Colorectal cancer is the second leading cause of cancer death in the western countries for both men and women [3]. The majority of affected colorectal cancer patients have sporadic disease with no

evidence of having inherited the disorder (Fig. 2). The estimated lifetime risk of developing sporadic colorectal cancer at the age of 50 is approximately 5% in the general population. Environmental and dietary factors are believed to increase this risk. Symptoms develop late in the course of the disease and approximately 40% of cases are diagnosed in late stages.

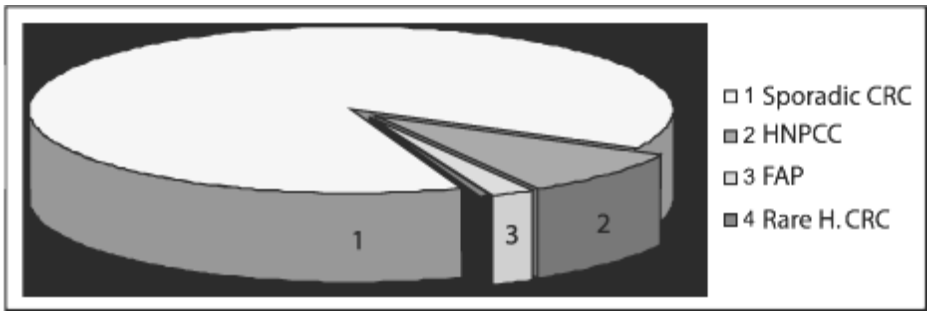

Figure 2. Type of Colorectal cancer (CRC). Sporadic and hereditary syndromes including hereditary nonpolyposis colorectal cancer (HNPCC), familial adenomatous polyposis (FAP) and rare hereditary colorectal cancer (H. CRC) such as Peutz Jeghers syndrome (PJS), juvenile polyposis syndrome (JPS) and mutY human homologue associated polyposis (MAP).

The genetic model of colorectal cancer is one of the best studied and is based on decades of research to identify the events that lead to tumorigenesis, and can provide a good model for other types of cancer caused by inactivation of tumor suppressor genes. Vogelstein et al. [4] proposed a multi-step model for the genetic events underlying the progression of colorectal cancer. In this model the transformation of adenoma to carcinoma is the result of a combination of genetic defects, including mutational events in oncogenes, tumor suppressor genes and stability genes that typically develop in a decade. Luebeck and Moolgavkar [5] schematically subdivided Vogelstein's multi-step model in four stages. The first stage is represented by normal cells. The inactivation of the APC gene is caused by two independent rare events which characterize the following two stages and result in a tumor initiation and adenomatous polyp. Development of thousands of adenomas form within 15-40 years is typical in familial adenomatous polyposis (FAP) patients. The step from the third to the fourth stage is designated by frequent genetic alteration in another gene, resulting in a clonal expansion of the mutated cells and malignant transformation. Several studies have shown that 95% of colorectal cancers have mutations in the APC pathway and additional mutations in genes of other pathways (e.g. p53 and K-ras). The four stage model concludes that pre-neoplastic events preceding the tumor involve several cell clones (polyclonal status), whereas the last stage leads to a selection and expansion of a single cell. Luebeck and Moolgavkar also hypothesized that after the third step cells of the colonic crypt loose the ability for asymmetric division and replicate

symmetrically, resulting in clonal expansion followed by the development of the adenomatous polyp.

A different model of colorectal tumorigenesis is supposed to characterize hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome. In this model inactivation of both alleles of the mismatch repair (MMR) genes results in a considerable increased rate of mutations in other genes (microsatellite instability) thus producing a mutator phenotype, facilitating the onset of cancer clones [6]. Colorectal microsatellite mutator phenotype (MMP) tumors do not have mutations in genes such as p53, but harbor genetic alterations in genes containing DNA repeats in the coding region, such as the transforming growth factor-β type II receptor (TGF-β RII) gene. These are the target for slippage-induced frameshift mutations during DNA replication that result in genomic instability. As a consequence genes other than MMR genes can play an important role in the progression from adenoma to carcinoma in HNPCC tumorigenesis. For example, mutations in the TGF-β receptor II gene may cause altered TGF-β type II receptor on the surface of the cell and an inability to bind transforming growth factor-β (TGF-β), followed by a decreased sensitivity to the growth inhibitory effects of TGF-β and tumorigenesis [7]. Perucho et al. [8] suggested a model based on "a mutator that mutates another mutator", where primary mutators such as *MLH1* and *MSH2* genes, cause frame-shift mutations in secondary mutators genes, such as *MSH3* and *MSH6* and thereby inactivating the MMR pathway and resulting in microsatellite instability.

The most frequent autosomal dominant hereditary colorectal cancer syndromes are HNPCC and familial adenomatous polyposis (FAP). Other inherited syndromes associated with increased colorectal cancer risk include juvenile polyposis, Peutz Jeghers and MAP (mutY human homologue associated polyposis, recessive condition) with very low incidences. Normally, an inherited colorectal cancer syndrome is identified by the following criteria: the presence of multiple affected individuals in the family of the patient, a young age at the time of diagnosis with the possibility of multiple cancers in the patient. Mutational screening is then performed in affected members of the family and at-risk relatives are tested to confirm the presence or absence of the inherited allele carrying the mutation. An inherited susceptibility to colorectal cancer involving tumor-suppressor and stability genes has been reported in less then 5% of all colorectal cancer cases (Fig. 2). Although these syndromes constitute only a small fraction, they are important because they provide a starting point to identify the genes and mechanisms involved in colorectal tumorigenesis. A more complete understanding of these syndromes may provide ways to detect and prevent carcinogenesis in at risk family members.

Some of the questions that will be addressed in this introduction and following chapters of this thesis are:

1. Can specific pathologic characteristics and pre-screening methods be a combined approach to select individuals affected by HNPCC and FAP eligible for mutational analysis?
2. Which are the genes responsible in colorectal tumorigenesis?
3. Which are the mechanisms leading to inactivation of tumor suppressor genes and mismatch repair genes?
4. Are some mechanisms more frequent in some genes than in others?
5. Are all mechanisms causing MMR gene silencing known?
6. Or are there unknown pathways that are less frequent and more difficult to identify?
7. Are direct mutation analysis and indirect haplotype analysis the only tests to be performed?
8. What are the limitations of the current tests?
9. Can a combination of multiple genetic tests improve the accuracy and efficiency of discovering genetic alterations in affected subjects?
10. Are new advance automated and fluorescent detection methods more rapid and efficient to enhance genetic screening on large sets of samples?
11. Which are the most powerful pre-symptomatic screening tests to detect subjects at risk of hereditary colorectal cancer?

Inherited Colorectal cancer syndromes

1.2 Hereditary Nonpolyposis Colorectal Cancer (HNPCC)

The first case of HNPCC was reported by Warthin at the end of the nineteenth century in a large family diagnosed with a variety of cancers [9-11]. Later analysis and documentation of this and other families by Lynch and colleagues in the 1970s and 1980s suggested that HNPCC is a common inherited cancer susceptibility syndrome. In its most striking examples HNPCC is characterized by a dominant mode of transmission with 80% penetrance (the remaining 20% of individuals with a predisposing mutation may never develop cancer) [12-16]. HNPCC accounts for 2-3% of reported colorectal carcinoma cases. The mean age of diagnosis is between 44 and 48 years compared to 64 years for sporadic colon cancers. The majority of tumors (60-80%) in HNPCC patients arise in the proximal colon (i.e. from cecum to the descending side). 10% of patients will have synchronous tumors at the time of diagnosis (i.e. a simultaneous onset of two or more distinct tumors in different parts of the bowel) or develop metachronous tumors (i.e. new tumors developing at least 6 months after the initial diagnosis). Histopathologic features of tumors show that the progression of HNPCC adenoma to adenocarcinoma is more rapid than that for sporadic colon adenomas and the carcinomas tend to be poorly differentiated with abundant mucin and lymphoid infiltration [17]. In the absence of screening most patients remain asymptomatic until the development of cancer.

In the early 1990s HNPCC was reported to be associated with multiple primary tumors and a high proportion of colorectal cancer as well as many other types of cancer, including endometrial, extracolonic gastrointestinal and ovarian cancer [18-21]. In particular, women with colorectal cancer develop endometrial cancer at an earlier age, with a life-time risk of 20-40%, compared to 3% in the general population [22, 23]. Ovarian cancer risk has been reported to be 3.5 times as high with a mean age of diagnosis 20 years earlier than sporadic cases [24]. Screening analyses such as colonoscopy can reduce the risk of colorectal cancer mortality by 50-65% [25]. International Collaborative Group-Hereditary Nonpolyposis Colorectal Cancer (ICG-HNPCC) guidelines recommend a full colonoscopy every one to two years beginning at the age of 20-25 for patients with a high risk for HNPCC (as determined by positive diagnostic criteria or genetic tests), as well as for first degree relatives of these patients [26]. Due to an associated risk for endometrial cancer, experts recommend annual screening by transvaginal ultrasound beginning at the age of 25-35 [27]. HNPCC in combination with benign or malignant sebaceous skin tumors is referred to as Muir-

Torre syndrome[28], whereas the development of glioblastoma multiform in association with HNPCC is called Turcot's syndrome [29]. Discussion of these latter syndromes lays outside the scope of this thesis.

Table 1. Clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC) [31-33]

Amsterdam I criteria	Amsterdam II criteria
1. At least three relatives with colorectal cancer (CRC), one of which must be a first-degree relative of the other two; 2. Colorectal cancer involving at least two successive generations 3. One or more colorectal cancer cases diagnosed before the age of 50 4. Familial adenomatous polyposis (FAP) has been excluded	1. Three or more relatives with HNPCC-associated cancer (colorectal, endometrial, small bowel, ureter, or renal pelvis), one of which must be a first-degree relative of the other two; 2. Colorectal cancer involving at least two successive generations 3. One or more colorectal cancer cases diagnosed before the age of 50 4. Familial adenomatous polyposis has been excluded

Bethesda guidelines	Revised Bethesda guidelines
1. Individuals with cancer in families who meet the Amsterdam criteria I-II 2. Individuals with two HNPCC-related cancer, including synchronous and metachronous colorectal cancers or associated extracolonic cancers (endometrial, ovarian,gastric, hepatobiliary, small bowel,or transitional cell carcinoma of the renal pelvis or ureter) 3. Individuals with CRC who have a first-degree relative with CRC and/or a HNPCC-related extracolonic cancer and/or colorectal adenoma; one of the cancers diagnosed before age 45 and the adenoma diagnosed before age 40 4. Colorectal cancer or endometrial cancer diagnosed before age 45 5. Individual with right-sided CRC, an undifferentiated pattern on histopathology diagnosed before age 45 6. Individuals with signet-ring-cell-type CRC diagnosed before age 45 7. Adenomas diagnosed before age 40	1. Colorectal cancer diagnosed before age 45 2. Presence of synchronous or metachronous colorectal or other HNPCC-associated tumors regardless of age 3. Colorectal cancer diagnosed before age 60 with histologic findings of infiltrating lymphocytes, Crohn's-like lymphocytic reaction, mucinous/signet ring differentiation or medullary growth pattern 4. Colorectal cancer diagnosed in one or more first-degree relatives with an HNPCC-related tumor, with one of the cancers diagnosed before age 50 5. Colorectal cancer diagnosed in two or more first or second-degree relatives with HNPCC- related tumors, regardless of age

The ICG-HNPCC adopted in 1991 a number of clinical criteria to standardize the study of HNPCC families. The original Amsterdam criteria (Amsterdam I) were focused to selected affected patients based on a family history of colorectal cancer and an early age of onset [30]. Due to the presence of multiple organ tumors in HNPCC families and exclusion of patients from small families, the

Amsterdam criteria were later revised to be more sensitive but less restrictive. In 1997 and 1999, Bethesda guidelines [31] and Amsterdam II [32] criteria were developed, respectively, to include individuals with a less well documented family history of cancer, the location of the tumor and the early age of the affected family member. These guidelines were recently again revised as the Revised Bethesda criteria by the ICG-HNPCC [33]. Thus, several different clinical criteria are used in an attempt to identify more accurately individuals who are likely to have a germline mutation in one of the MMR genes (Table 1).

1.2.1 Molecular genetics

1.2.1.1 MMR complexes and repair activity

Bacterial and eukaryotic cells have a mismatch repair (MMR) system that recognizes and repairs single-base mispairs that arise during DNA replication, as well as small insertions and deletions. Cells with a deficient mismatch repair system have mutation rates 100 to 1000 times as high as normal cells, often resulting in system replication slippage in homopolymeric repetitive sequences, known as microsatellites. The mismatch repair system was originally identified in bacteria where the homodimers MutS, MutL and MutH encoded by MMR genes, operate soon after incorporation of a wrong base just before the newly synthesized daughter strand becomes methylated [34-36]. The discovery of MMR genes in yeast helped to identify the homologous proteins in human, as the MMR system is highly conserved from bacteria to mammals (Fig. 3). The MMR system became more important with the discovery of the inherited deficiency of the corresponding human genes and the association of these with HNPCC. The human MMR system is probably more complex the bacterial system. The system involves several proteins encoded by a number of genes, including *MLH1*, *MSH2*, *MSH6*, *PMS1*, *PMS2*, *MLH3* and *MSH3*, but only some of them are known to cause hereditary nonpolyposis syndrome (Table2) [37, 38, 39].

Table 2. Molecular characteristics of the mismatch repair genes associated with HNPCC

Human	Chromosomal location	Mutation rate in HNPCC (%)	E. coli homol.	Yeast	cDNA size	Genomic structure (n. of exons)	Amino acid	References
MSH2	2p21-22	30	MutS	yMSH2	2804	16	934	39
MSH6	2p16	6	MutS	yMSH6	4082	10	1360	38
MLH1	3p21-23	63	MutL	yMLH1/yPMS2	2268	19	756	37
PMS2	7p22	0.4	MutL	yPMS1	2588	15	862	37

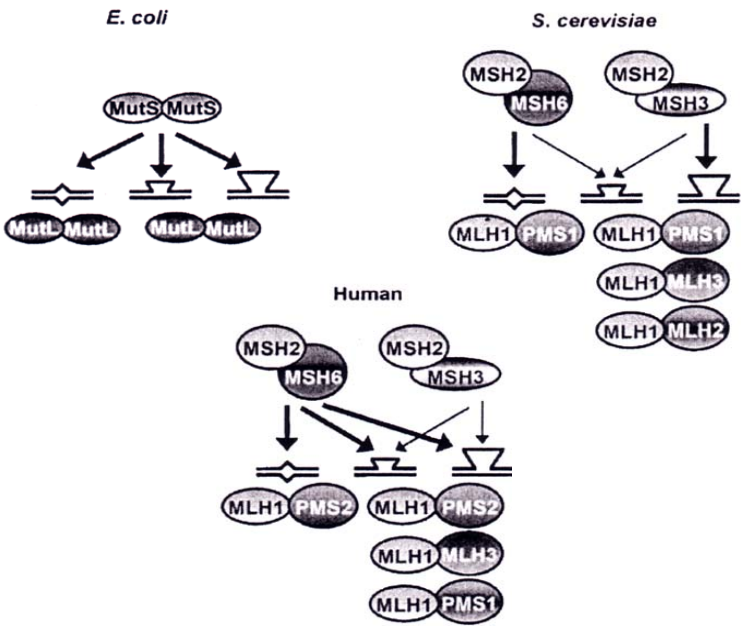

Figure 3. Schematic representation of the mismatch repair system homologues in different organisms. Base-base mismatches, +1 insertion/deletion (IDLs), and larger loops with 2 to about 8 unpaired nucleotides (2-4) in E. coli) (shown for each organism form left to right) are bound by MutS and eukaryotic MSH heterocomplexes. Thick arrows indicate specificity to the DNA heterologies. Thin arrows indicate a function as backup or redundancy system between MSH2-MSH6 and MSH2-MSH3. MutL homologues involved in subsequent repair of either base-base mismatches or loops are given below the DNA heterologies. Adapted from Marti et al., J. Cell. Physiol. 2002; 191:28-41. Reproduced with permission from John Wiley & Sons, Inc.

During the synthesis of a new DNA strand some of the replication errors are immediately corrected by the 3' to 5' exonuclease activity of DNA polymerase, which reduces the rate of replication error to approximately 1 per 10^{12} base pairs. The mutations that escape the proofreading DNA polymerase activity are repaired by the MMR system. Genetic and biochemical data [40-43] showed that in human and in yeast DNA strands carrying a mismatch are recognized and bound by a complex consisting of *MSH2* protein and either *MSH6* or *MSH3* proteins. The *MSH2-MSH6* heteroduplex, called MutSα [41], is able to bind most base/base mismatches and small insertion/deletion loops (IDLs), whereas the *MSH2-MSH3* heterodimer, called MutSβ [43], preferentially binds to small and large IDLs. After that MutSα/β complex binds to a DNA mismatch, the identification of the newly synthesized strand is achieved by the MutL complex, which binds to the DNA mispair-MutSα/β complex (Fig. 4). In cells with normal MutL function, the MLH1 protein forms three different heterodimeric complexes with *PMS2*, *PMS1* and *MLH3* proteins. All of these proteins interact with the same region of the *MLH1* protein, but the first complex (*MLH1-PMS2*) is more abundant then the latter two heteroduplexes (*MLH1-PMS1*; *MLH1-MLH3*) and provides the majority of repair activity [44]. The repair system also requires ATP,

the binding of which induces a conformational switch of MutS into a sliding clamp that is hypothesized to leave the mismatch and move along the DNA to search for downstream MMR factors [45, 46].

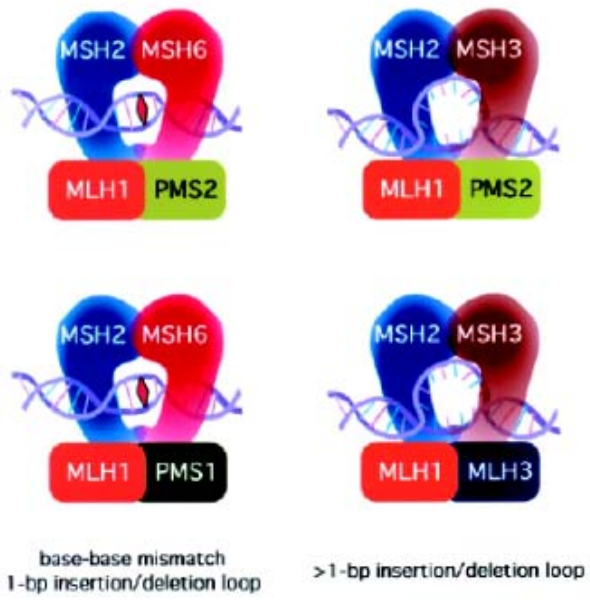

Figure 4. Functional interaction of hMutL and hMutS heterodimeric complexes involved in the repair of base-base mismatches (red diamond) and insertion/deletion loops. MutSα (*MSH2-MSH6*) and MutSß (*MSH2-MSH3*) are represented with general shape of two "opposing comas", forming a channels occupied by the mismatches DNA. This MutS model includes interaction with MutL on one side and ATP-binding site on the opposite side (not shown). Adapted from Bellacosa A., Cell Death Differ. 2001;8:1076-92. Reproduced with permission from Nature Publishing Group.

There are additional components in the human MMR system, including helicase II, human 5'>3' exonuclease EXO I, human polymerase δ, PCNA, RPA and DNA ligase I [47-49]. The function of the MMR system is to excise a DNA fragment containing the mismatch and fully resynthesizes the correct sequence. The mismatch recognition is mainly performed by the MutSα heterodimer, which contains about 90% of MSH2 and is present at much higher levels than MutSβ in the cells, suggesting that MutSα has a primary function in the MMR system [50]. Interestingly, some authors reported that inactivation of MHS6 leads to a loss in repair of single-base mismatches and an instability of mononucleotide repeats, whereas a lack of MSH3 appears to have a weak effect [51, 52]. However, in the absence of MSH6 and the consequent MutSα, the partial redundancy function of MutSβ mediates the repair of insertions/deletions larger than 2 basepairs and MSH3 is also found to be present at high levels in cells [42]. Several groups showed that a lack of hMSH2 or hMLH1 protein results in a concomitant absence of the heterodimeric partners and in the instability of the hMutS or hMutL complexes even in the presence of normal *MSH6*, *MSH3* or *PMS2* proteins [50, 52-54]. Mutations of *MLH1* and *MSH2* genes have been suggested to cause an early onset of colorectal cancer [55]. Human MSH2 also interacts with two other genes, *MSH4* and *MSH5*, to form heterodimers that are involved in meiotic recombination [56].

1.2.1.2 The two-hit model

MMR genes are believed to behave like tumor-suppressor genes. According to the two-hit model of carcinogenesis (the Knudson model) two mutational events are required to cause a loss of function [1, 2]. In HNPCC patients, constitutional cells carry germline alterations in one copy of the MMR gene (first hit) and one normal copy (wild-type). Normally, the MMR gene is functional until a spontaneous somatic event creates a defect in the wild-type allele (second hit), and silences the gene. Interestingly, four different pathways have been proposed as a first hit mechanism inactivating MMR genes (Fig. 5).

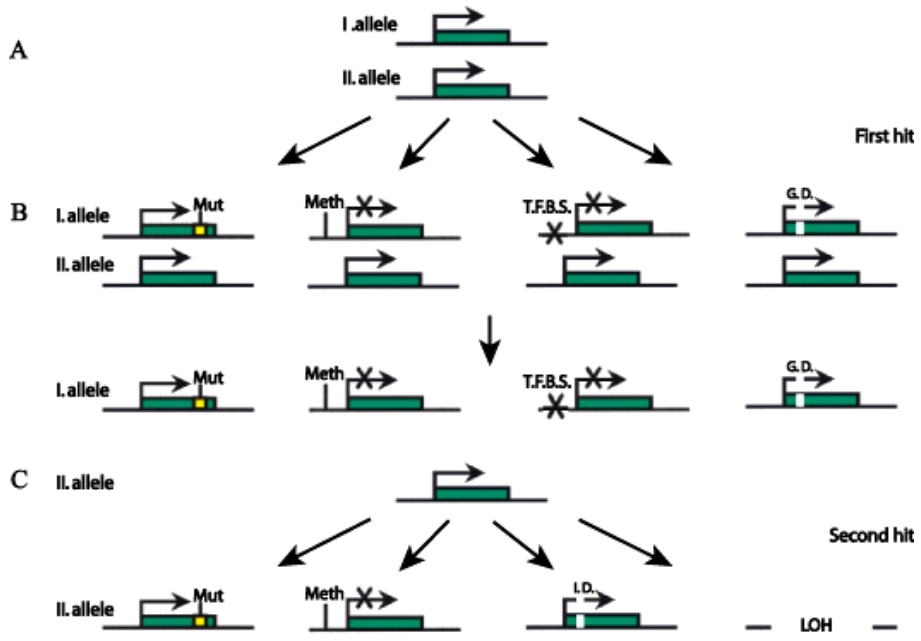

Figure 5. Knudson's two-hit hypothesis revised. Schematic representation of the mechanisms involved in the allelic inactivation of tumor suppressor and mismatch repair genes. **A**. Allele I and II are indicated by two green boxes shown at the top (arrow) transcriptional activity of the gene in normal cells. **B**. Four different genetic alterations (mutation, promoter methylation, altered transcription factor binding site and genomic deletion) can cause inactivation of the allele I, representing the first hit at germline level which results in the predisposition to the hereditary form of colorectal cancer. **C**. During the life-time additional genetic alterations (second hit) lead to the inactivation of the allele II in the somatic cells, causing full transcriptional silencing of the gene. (Mut.-mutation; Meth.-promoter methylation; T.F.B.S.-transcription factor binding site; G.D.-genomic deletion; I.D.-intragenic deletion and LOH- loss of heterozygosity). Adapted from Jones P.A. and Laird P.W., Nat Genet. 1999 Feb; 21(2):163-7. Reproduced with permission from Nature Publishing Group.

In HNPCC patients, the first pathway involves mutations in the *MLH1* or *MSH2* genes in the majority of cases and rarely in *MSH6* and *PMS2* genes. These genetic alterations include missense

or nonsense mutations or insertions or deletions of base pairs, and it has established as the most common first step of allele inactivation at the germline level. Recently, new techniques have been developed to detect rearrangements of intra or whole genes which identified genomic deletions as a second pathway of single allele inactivation in MMR genes. Some investigators have suggested epigenetic changes in the promoter as a first hit at the germline level as a third alternative potential pathway (**Chapter 5**, [57-60] [61]). A fourth possible mechanism involves putative regulatory mutations in the promoter of *MLH1*, *MSH2* and *MSH6* with effects on the MMR system. Polymorphisms in the transcription factor binding site may decrease the affinity for the nuclear factor resulting in a reduced promoter activity and expression of the gene [62-64]. An example of this is reported in **Chapter 6.** It is worth noting that the latter two mechanisms result in a possible decrease of mRNA expression with no change in the DNA sequence. In contrast, genetic alterations in the first two mechanisms (like mutations or genomic deletions of partial/total genes) affect the coding sequence with a possible effect on the structure of the protein. At a somatic level (the second hit) in HNPCC and colorectal cancer patients inactivation of the second allele is caused by point mutations, insertions and deletions of a base pair, hypermethylation or loss of heterozygosity (LOH). In this thesis the latter two pathways are described in two different sets of HNPCC patients providing new mechanisms of developing susceptibility for colorectal cancer.

1.2.2 Genomic instability

A deficiency of MSH2, MLH1 and other MMR genes results in many DNA replication errors, usually in repetitive sequences of one to a few base pairs that are particularly prone to polymerase slippage. These simple sequence repeats (or microsatellite repeats) occur both in the regulatory and the coding regions of many genes. There are 50.000-100.000 of these mono-dinucleotide repeat sequences in the genome of which the most common are (A)n/(T)n and (CA)n/(GT)n. Such errors lead to an effect called microsatellite instability (MSI) or replication error positive (RER+) phenotype, which is only detected in tumor tissue. From individual to individual the number of repeats can be variable but in the same person the microsatellite sequences are fixed in all normal cells. They are particularly important in linkage analysis, genomic mapping and also in the detection of loss of heterozygosity (LOH) in tumor suppressor genes. Usually microsatellite instability (gain or loss in the number of repeats in a DNA sequence) is identified by comparing the length of the microsatellite in normal tissue and tumor tissue in the same individual. The length of the repeated sequences is measured by performing polymerase chain reaction (PCR) using primers

that flank these regions. An example of microsatellite instability in stable and unstable tumor cells in the same individual is given in figure 6. Microsatellite instability consists of new alleles corresponding to a shorter or longer stretch of repeats resulting in an electrophoretic mobility shift of the peaks in tumor compared to normal tissue. Genomic instability is thought to be responsible for the rapid accumulation of somatic mutations in cancer-related genes, such as oncogenes and tumor suppressor genes. Genes that posses repeats in their coding region appear to be an easy target for frameshift mutations by incorrect replication, resulting in high genetic alteration rates in tumor cells with a defective MMR system, also known as microsatellite mutator phenotype (MMP) [8, 65]. Some authors have reported microsatellite instability in several genes, including *TGFβ-RII* [66], *APC* [67], *BAX* [68], *MSH3*, *MSH6* [69], *BRCA1*, *BRCA2* [70], *PTEN* [71] and the transcription factor *E2F4* [72]. The link between MSI and MMR deficiency was first identified in HNPCC tumors [73-75] and later in many others, including sporadic colon [76, 77], lung [78], breast [79], stomach [80], pancreatic [81], uterine [82], and endometrial and ovarian [83, 84] cancers. Microsatellite instability has been identified in more than 90% of HNPCC tumors compared to only 10-15% of sporadic tumors [75, 77, 85]. Colorectal tumors with MSI are diploid and generally located in the ascending or the transverse colon [74, 86].

1.2.2.1 Microsatellite instability guidelines

The MSI status was proposed by the National Cancer Institute (NCI) Workshop on Microsatellite Instability for Colorectal Cancer Detection [87], which was approved and comprised in the Bethesda guidelines [31] and later revised in 2002 [33]. According to the HNPCC-NCI guidelines a panel of five microsatellite loci is recommended to detect with sensitivity and specificity the MSI in HNPCC patients that have potential mutations in the MMR genes. This panel (Fig. 6) contains two mononucleotide repeats (BAT 25 and BAT26, polyA) and three dinucleotide repeats (D2S123, D5S346 and D17S250), where tri-, tetra- and pentanucleotide microsatellites are not included because they appear stable. However, the addition of more mono-dinucleotide markers is suggested in case where the evaluation of MSI is inadequate. Following the NCI recommendations, tumors are classified as MSI-High (MSI-H) when two or more of the five microsatellite sequences are mutated compared to normal tissue from the same patient, and in MSI-Low (MSI-L) or MMS if only one or none of the markers show instability (Fig. 6). MSI analysis has become an important diagnostic tool for identifying those HNPCC cases that are due to MMR defects. A high proportion of Amsterdam criteria HNPCC cases show tumors with microsatellite instability. Through the application of these guidelines most HNPCC cases that have MSI-H tumors have been linked to a mutation in MSH2 or MLH1 genes.

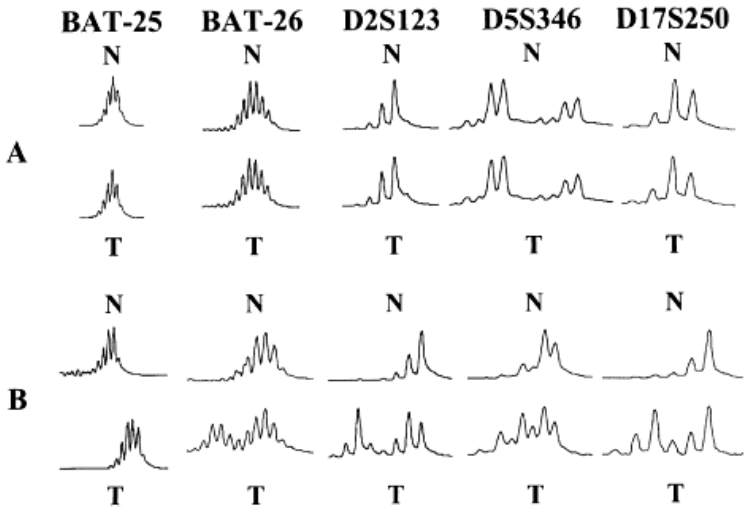

Figure 6. Representative examples of microsatellite instability at the five locus markers. Peaks visualized on the ABI377 sequencer and corresponding to normal and tumor DNA are labelled N and T, respectively. **A** and **B** are examples of stable and unstable microsatellites in colorectal cancer cases, respectively. Adapted from Colombino M. et al., Ann Oncol. 2003 Oct;14(10): 1530-6. Reproduced with permission of Oxford University Press.

A further proof that mutations in *MLH1* and *MSH2* are associated with MSI came from a study of Umar et al. [18], who reported a correction of MMR deficiency (a homozygous mutation in the MLH1 gene) and microsatellite instability by introducing a single chromosome 3 from normal fibroblasts in a human colon tumor cell line (HCT116). As is reported for the MLH1 gene [18], the introduction of a wild-type MSH2 gene into a MSH2-deficient tumor cell line reduces the accumulation of mutations to normal levels [88]. In the mouse a lack of MSH2 or MLH1 genes results in a loss of MutS or MutL complexes and an incomplete inactivation of the MMR system, which increases mutations in all tissues examined [89, 90]. In contrast, a deficiency of the others proteins involved in the complexes, results in less restricted effects due to the redundancy function between MutS α and MutS β [91, 92].

A recent study reported that 95% of HNPCC cases with MSI-H show a related loss of *MLH1* (40%), *MSH2* (40%), *MSH6* (10%) and *PMS2* genes (5%) [93]. In addition, in a few cases where inherited mutations were not found, an altered expression of MSH2 or MLH1 was identified suggesting the presence of defect in these genes. According to the international guidelines, genotyping of the tumor for MSI and immunohistochemical analysis should be the primary screening methods after verified positive family history.

1.2.3 Genetic predisposition

The discovery of the role of MMR genes in HNPCC has led to the identification of many families with inactivating germline mutations. The development of novel techniques, described below along with their advantages and disadvantages, has considerably simplified the screening of DNA mismatch repair genes. The heterogeneity of the identified mutations suggests an absence of correlations between genotype and phenotype, and a lack of mutational "hotspots" in the coding sequences results in a long and difficult screening process, which requires the full sequencing of the genes. At the moment no guidelines exist for classifying genetic alterations as pathogenic mutations or polymorphisms, resulting in a great variation in interpretation by different researchers.

1.2.3.1 Inactivation of MLH1 and MSH2 genes by mutations

In the last decade the investigation of HNPCC families has shown that most germline mutations are in the *MLH1* and *MSH2* genes [94-97], which are the major cause of DNA-mismatch-repair deficiency in humans. Frameshift and nonsense mutations (protein truncating mutations), missense mutations (protein sequence changes) and deletion mutations have all been observed and reported in a database (http://www.insight-group.org) [98]. In this database more than 448 different predisposing mutations are described, of which 50% affect *MLH1*, 39% *MSH2*, 7% *MSH6* and the remaining genes are together responsible for less then 4%. The germline mutations that occur in *MLH1* and *MSH2* appear widely distributed throughout either genes, with a prevalence of frameshift and premature truncations in *MHS2* (60% and 23% of all mutations, respectively) and missense mutations in *MLH1* (31% of total mutations) [99]. In the last few years several authors reported a high percentage of genomic deletions in *MLH1* and *MSH2* genes [100-104], whereas germline epigenetic alterations are mainly observed in *MLH1* [57-59] and only in one case in *MSH2*[61]. Several recurrent and "founder germline mutations" have been observed in MLH1 and MSH2 genes, which appear to share a common ancestor. These founder mutations have been identified in the Finnish population [105, 106], where two genetic alterations in *MLH1* (3.5-kb in exon 16; splice-site mutation at exon 6) account for the majority of reported families. In America a founder mutation in the *MLH1* gene has been observed in Navajo families from Arizona, and a founder mutation in *MSH2* (A636P) has been identified in Ashkenazi Jews [107-109], which arose approximately 17 generations before during the time when this community were living in eastern Europe. Founder mutation has been also reported in *MSH2* (splice donor exon 5) to affect the English and North American HNPCC families [110]. Using Southern blot and DGGE, Wagner et al [111] identified a large founder deletion of 20-kb in

MSH2 (exon1-6). Interestingly, genealogical, clinical and molecular studies traced this mutation in 14 different States in U.S., confirming that was a common North American founder deletion. In addition, the origin of the mutation could be traced back to German ancestors in the early 18th century [112].

1.2.3.2 MSH6 and its important role

Many other genes encoding proteins involved in the MMR system have been considered as potential mutator genes implicated in the tumorigenesis of HNPCC. Of these, the best studied is MSH6. Defects in the MSH6 gene cause only a partial loss of MMR activity, as MSH3 can compensate for defects in MSH6. Mutations in the MSH6 gene result in a large increase of base substitution mutations. Germline mutations in *MSH*6 have been identified in a small number of HNPCC families and in a large number of atypical-HNPCC cases with late onset [113-119]. Dominant mutations have been reported at the N- and C- terminals of the MSH6 protein, where both the MSH2 interaction domain and the ATP-binding site occur [116, 118, 120]. The MSI status of the cases with MSH6 mutations is not fully clarified. Some studies characterize the tumors as MSI-L, whereas others suggest a MSI-H phenotype. A MSI-H phenotype may be due to a combination of *MSH3* and *MSH6* mutations. However, human cell lines with *MSH6* mutations show MSI in mononucleotide and not in dinucleotide repeats [121]. Studies have reported that HNPCC families with germline *MSH6* defects have a high proportion of endometrial cancer [119, 122]. In contrast, approximately 8% of familial colorectal cancer (non-HNPCC) cases were found to have germline *MSH6* defects. Similarly, germline *MSH6* defects have been found in endometrial cancer cases associated with a weak family history of cancer. Patients with mutation in *MSH6* gene appear to have less striking family history and later age of onset of cancer compared to HNPCC cases with mutation in *MSH2* or *MLH1*. A correlation was observed between the expression of MSH2 and MSH6 proteins in HNPCC and endometrial cancer cases. Immunohistochemical analysis showed a reduction of the expression of MSH2 protein in the presence of germline mutations in the MSH6 gene, whereas a lack of MSH6 protein was observed in cases with germline mutations in the MSH2 gene [54, 123]. It has been observed that mouse cells that are homozygous for MSH6 mutations develop tumors at late ages with a reduced expression of MSH2 and low or no microsatellite instability [124, 125]. Recently, few studies reported genomic deletions in MSH6 gene in HNPCC and in suspected HNPCC families [104], [126] While more work on *MSH6* is needed, current evidence suggests that germline *MSH6* defects can cause an attenuated form of HNPCC and inactivation of *MSH6* causes a weaker mutator phenotype confined to base substitutions and insertions/deletions in the mononucleotide repeats [127, 128].

1.2.3.3 Others MMR genes

A number of other MMR genes have been considered to be involved in HNPCC, including *PMS1*, *MSH3*, *MLH3* and *PMS2*. However, defects in some of these genes have not been found to make major contributions to HNPCC. This could be consistent with studies indicating that defects in these genes do not result in a complete loss of the MMR system because they encode partially redundant functions. There is no clarified correlation between mutations in the *PMS1* gene and HNPCC. One study described a germline mutation in the *PMS1* gene [129] in HNPCC patient, but no other cases have been described since. However, the original *PMS1* mutant family was later shown to contain an additional *MSH2* mutation [130]. Mutations in the *MLH3* gene are also likely to be rare in HNPCC. Wu et al. [131] reported ten germline mutations including one frameshift and nine missense mutations in twelve suspected HNPCC patients. Only in few of them the identified mutations could be associated with the development of colorectal cancer and thus be considered putative pathogenic alterations. One study identified potential missense mutations but these variants could be rare polymorphisms [132]. Most subsequent studies have not identified germline *MLH3* mutations in HNPCC cases. Recently, Liu et al. [133] screened 70 families with hereditary colorectal cancer both high and low risk and identified one frameshift and eleven missense mutations, but no microsatellite instability was detected. Three mutations showed segregation with the disease and most were identified in families with low risk, containing one or two affected first-degree relatives. Interestingly, the authors reported one case, in which the paternal allele carried a *MSH*2 missense mutation and the maternal allele presented a *MLH3* variant, suggesting that two combined mutations can caused stronger mutator phenotype than a single mutant. This idea is supported by a previous study on the functional interaction between weak alleles in yeast and the resulting mutator phenotype [134]. No germline defects in the *MSH3* gene have been reported. It is conceivable that defects in these genes could cause cancer susceptibility in late onset and in families with a weaker history than usual in HNPCC. In contrast, 7 truncating mutations in PMS2 gene have been reported in 7 different HNPCC suspected families [129, 135-137], including 3 fulfilled the Bethesda criteria and 1 the revised Amsterdam criteria. Inactivating mutations in this gene have also been observed in affected individuals with Turcot's syndrome [138, 139]. Recently, Hendriks et al. [140] have identified 4 genomic rearrangements and 3 truncating point mutations in PMS2 gene in HNPCC suspected families, which fulfilled Amsterdam criteria. All 7 tumors were MSI-H and showed loss of the expression of the PMS2 protein. Germline genetic alterations in the *PMS2* gene appear only in few cases of HNPCC families and not always fulfill Amsterdam criteria. This attenuated phenotype can be explained by the redundancy function with MLH3, which can interact with MLH1 in the

absence of PMS2. Moreover, identification of genetic alterations may be difficult due to the presence of a pseudogene close to PMS2 gene.

1.2.3.4 Epigenetic effects

Alterations in the methylation patterns of the promoter region of tumor suppressor genes have been reported to be common event in human neoplasia and have been associated with a decreased expression of the genes [141, 142]. Epigenetic alterations are characterized by methylation at the cytosine residues of CpG dinucleotides, which does not change the DNA sequence. Sequences rich in CpG (CpG islands) are often located near the 5'terminus of the gene, which contain the promoter, the 5' untranslated region and the first exon. Methylation of the CpG islands in the promoter may result in silencing of the gene [143]. The mechanism by which hypermethylation represses the transcription of the gene is not clear, but it has been suggested that methylation per se blocks the binding of transcription factors (see **Chapter 6**) or that methylation modifies the chromatin structure thereby inhibiting access of the transcription factors. In normal human cells CpG islands are usually unmethylated. Not all methylated CpG sites inactivate the gene. It has been reported that hypermethylation of exonic CpG islands may occur without effect on the expression of the gene itself [143]. In many sporadic colon cancers, loss of MLH1 expression is due to somatic silencing of both copies of the *MLH1* gene associated with biallelic methylation or methylation and LOH or mutation and methylation [144-147]. Some studies also showed that methylation of the promoter and absence of the MLH1 protein are associated with MSI phenotype colorectal cancer [144, 148, 149]. This correlation has also been observed in sporadic endometrial cancer associated with HNPCC, where the proportion of cases showing MSI phenotype with identified hypermethylation in the *MLH1* promoter, or mutations in the *MLH1* or *MSH2* genes is 20% to 34% [150-154]. Promoter methylation of multiple genes has been reported in non-malignant gastric tissue [155, 156], non neoplastic prostate tissue [157] and ulcerative colitis [158]. Epigenetic effects have also been suggested to be related to age [156, 158] or malignant stage [159] and in some cases associated with environmental risk exposures [160]. The detection of epigenetic abnormalities using classical methods, based on digestion with restriction endonucleases (Hpa II and MspI) sensitive to methylated or unmethylated CpG sites followed by amplification, has been replaced by more sensitive techniques such as methylation-specific polymerase chain reaction (MSP). In the MSP assay genomic DNA is modified by using sodium bisulfite, which converts all unmethylated cytosine residues to uracil but leaves methylated cytosines unchanged. To confirm that the modification has been successful, the sample is amplified with primers specific for the unmethylated sequence, or the methylated sequence, or sequences that do not contain CpG sites. Subsequent sequence analysis determines the

presence and the specific locations of the methylated CpGs. An example is given in **Chapter 5**. The use of MSP analysis of the methylation status of the *MLH1* promoter recently showed epigenetic alterations in peripheral blood from HNPCC or early sporadic colorectal cancers, suggesting a new mechanism of inactivation of one allele [57-60]. Recently, a first case of transgenerational inheritance in human has been reported by Chan et al. [61] The authors identified germline hypermethylation in the *MSH2* promoter in ten members of a putative HNPCC family with MSI, using allelic specific methylation assay and haplotype analysis. The segregation of hemiallelic methylation of *MSH2* in the family members provided strong evidence for an inherited germline epimutation.

1.3. Familial adenomatous polyposis (FAP)

The first case of polyposis was reported in 1721 [161], but only in the late 19th century polyps in the lower bowel were associated with inherited familial cancer [162 163, 164]. Today familial adenomatous polyposis (FAP) is recognized as a rare autosomal dominant syndrome of the colon and is characterized by the development of hundreds to thousands of intestinal adenomatous polyps at an early age. The syndrome affects both sexes equally and is responsible for approximately 1% of all colorectal cancers. FAP arises in affected individuals during their teenage years. Patients that carry identified mutations should begin colonoscopy exams at the age of 10-12 years, whereas individuals at risk for FAP with unknown genetic alterations should have periodical screening colonoscopies by the age of 15, annually from 26-35 and every three years from the age 36-50 [165]. If colectomy is not performed after the polyps emerge, the development of multiple adenomas through all the colorectum is inevitable, which it is thought to occur approximately 10-15 years after the onset of polyposis [166]. The average age at diagnosis of FAP ranges from 34-43 years [166, 167] which is early compared to the sporadic colorectal adenoma (around the age of 70). The first symptoms are blood per rectum, vague abdominal pain, and diarrhea with or without obstruction. When FAP is suspected the number and the nature of the polyps need to be verified. Affected FAP patients can also develop a variety of extracolonic malignancies in the upper gastrointestinal [168]. In addition, several extracolonic manifestations outside the gastrointestinal tract can occur with increased frequency in patients with FAP, which result in different clinical features but with defects in the same adenomatous polyposis coli (*APC*) gene. Three different variants have been reported to be associated with FAP. The first one is characterized by multiple colonic adenomatous polyposis, congenital hypertrophies of the retinal epithelium (CHRPE), osteomas of the skull and mandible,

epidermoid cysts, fibromas, desmoid tumors (Gardner syndrome) [169] [170-173]. The second one is characterized by the association of colonic adenomatous polyposis with the development of brain tumors (medulloblastomas), (Turcot variant) [174-176]. This syndrome can be a genetically heterogeneous condition, as in some cases, it has been reported to be associated with HNPCC in the presence of glioblastomas of the CNS [175, 177]. The third one is characterized by the presence of less than 100 polyps, which are often flat adenomas in the right-side of the colon. Cancer is mostly diagnosed in patients in the late onset (Atypical FAP) [178].

1.3.1 Genetic defect

In 1987 two groups reported a chromosomal deletion in affected FAP patients using genetic linkage analysis [179, 180]. A few years later the adenomatous polyposis coli (APC) gene was localized and characterized using positional cloning [181, 182]. The APC gene was identified to be associated with the FAP syndrome. The APC gene is a tumor suppressor gene composed of 15 exons and contains an open reading frame of 8529 bp that encodes for a 2843 amino acid protein with a molecular weight of 312 kDa. Almost all cases of FAP are caused by germline mutations of the APC gene on chromosome 5q21-22. The APC protein is considered important in cell adhesion, signal transduction and transcriptional activation. More than 1000 different germline mutations in the APC gene have been published and later assembled in an APC mutation database [183]. Ninety four % of APC germline mutations lead to a truncated protein product: 33% of APC mutations are nonsense point mutations, 6% small insertions and 55% small deletions. The frequency of detected APC mutations varies with the technique used (SSCP, DGGE, DNA sequencing). Because of the high number of premature stop codons, the *in vitro* synthesized protein assay (IVSP) or also known as the Protein Truncation Test (PTT) is still the most commonly used method. The inability to identify all mutations has required the development of a variety of new assays described in the appendix that increase the sensitivity to 90-95%. According to different studies, genotype-phenotype correlations have been associated to different clinical manifestations, suggesting that the position of *APC* mutations may confer a different development and risk of the disease [184, 185]. Germline mutations located between codons 169 and 1400 result in classic FAP, while the proximal or distal parts of the *APC* gene predispose to a milder phenotype called Atypical FAP [186-188]. An increased risk of severe colonic polyposis has been reported in the mutation cluster region (MCR) between codons 1250 and 1400 [189, 190]. Correlations between the position of the inherited mutation and the location of the somatic mutation have been observed. In fact, combinations of germline

mutations outside the MCR are usually followed by a second mutation within the MCR (1286-1513). However, inherited mutations between the codons 1100 and 1400 are typically associated with LOH [191, 192 193]. Moreover, predilection for extracolonic manifestations, such as desmoid or mandibular lesions, is associated with mutations identified between codons 1403 and 1578 [194], and congenital hypertrophy of the retinal pigment epithelium (CHRPE) with mutations in the region between codons 463 and 1387 [195, 196]. Deletions of 5 base pairs at codon 1309 and at 1061 are the most frequent mutation in FAP patients, (in respectively 18% and 12% of the patients) although mutations in the 3' region of *APC* gene are rarely detected [185]. A common polymorphism (I1307K) was identified in 6% of Ashkenazi Jews and 28% of Ashkenazi with a family history of colorectal cancer, resulting in a twofold increased lifetime risk of colorectal cancer [67, 197]. This polymorphism is a T-to-A transversion, which occurs within a stretch of adenine repeats (AAATAAA) in the APC gene, resulting in a sequence that is prone to polymerase slippage. Interestingly, analysis of the tumors showed an increased frequency of somatic mutations in the allele carrying the polymorphism. This genetic alteration (and also the one reported in **Chapter 6**) shows that polymorphisms can predispose to additional genetic alterations and increases the risk of colorectal cancer [198].

In the last few years, the development of fluorescence technology has improved the identification of genetic alterations and the sensitivity of the analysis. Different authors indicate the combined use of different techniques and direct sequencing as an efficient strategy for the detection of mutations in FAP and HNPCC. In addition, advance technology increases the sensitivity of genetic testing, resulting in a decrease of mortality due to a better treatment and follow up in patients carrying genetic alterations.

1.4 Outline of the thesis

As described above, transcriptional silencing of tumor suppressor genes and MMR genes requires the inactivation of both maternal and paternal alleles (Knudson's two-hit hypothesis). In the hereditary forms of colorectal cancer, the first allelic inactivation is inherited and the second is caused by a somatic mutation. A variety of point mutations have been identified that inactivate the germline allele of MMR genes. However, mutations have not been identified in a considerable number of HNPCC cases due the presence of rare inactivating mechanisms or genetic alterations in

unknown genes or to the use of few sensitive techniques. The aim of the research described in this thesis was to elucidate alternative mechanisms of germline genetic alterations involved in the first step of hereditary colorectal carcinogenesis.

In **Chapter 2** a linkage analysis was performed on 20 FAP families with no identified point mutations in the *APC* gene to identify pre-symptomatic individuals. The data showed that linkage analysis still has a high rate of success in defining the risk of presymptomatic subjects in FAP cases where mutational analysis failed. The results indicated that the assay is sensitive and the combination of direct mutation and linkage analysis could improve diagnostic screening. In **Chapter 3** HNPCC and suspected hereditary colorectal cancer families were analyzed for MSI, to verify whether this assay represents a tool to manage identification of suitable cases for further mutational analysis in MMR genes. We found that MSI was associated with 85% of HNPCC cases. New and others previously identified mutations (Chapter 4) were detected in HNPCC families with MSI-H. The results suggested MSI analysis as a good predictive test to identify HNPCC cases that are likely to be mutation carriers. In **Chapter 4** mutational analyses was performed in 16 Italian HNPCC families, in order to identify germline allelic inactivation of *MLH1*, *MSH2* and *MHS6* genes and to facilitate presymptomatic diagnosis of HNPCC family members. Several mutations were identified in *MLH1* and *MSH2* genes in germline cells, even though the sensitivity of the analysis was limited by the use of few techniques. The results showed that the mean age of tumor onset in HNPCC families with identified mutations was earlier then in patients without mutations, suggesting that the identified genetic alterations could cause a more severe phenotype. The second part of this thesis describes new mechanisms of MMR gene inactivation. In **Chapter 5** a first case is described in which constitutional hemi-methylation in the promoter of the *MLH1* gene causes transcriptional silencing. Additional LOH at the somatic level resulted in a complete absence of MLH1 protein in the tumor tissue. This finding showed a new and rare mechanism inactivating one allele of *MLH1* gene in germline cells. **Chapter 6** reports the characterization of the *MSH6* promoter region and its regulation by seven Sp1 transcription factor binding sites. Genetic alterations were observed in these transcription factor sites that caused a partial inactivation of the *MSH6* gene. Data showed that transcriptional activity of *MSH6* gene, containing these genetic alterations, was 50% lower then the wild-type. In addition, sixteen percent of the Caucasian population was identified to have these polymorphic variants, suggesting that human population contains individuals with different levels of *MSH6* promoter activity. **Chapter 7** presents a general discussion of the results described in this thesis.

References

1. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971;68:820-3.
2. Knudson AG, Jr. Hereditary cancer, oncogenes, and antioncogenes. Cancer Res 1985;45:1437-43.
3. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.
4. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-32.
5. Luebeck EG, Moolgavkar SH. Multistage carcinogenesis and the incidence of colorectal cancer. Proc Natl Acad Sci U S A 2002;99:15095-100.
6. Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res 1991;51:3075-9.
7. 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.
8. Perucho M. Microsatellite instability: the mutator that mutates the other mutator. Nat Med 1996;2:630-1.
9. Warthin A. Heredity with reference to carcinoma. Arch Intern Med 1913;12:546.
10. Warthin A. The further study of a cancer family. J Cancer Res 1925;9:279.
11. Warthin A. Heredity carcinoma in man. ANn Intern Med 1931;4:681.
12. Lynch HT, Krush AJ. Cancer family "G" revisited: 1895-1970. Cancer 1971;27:1505-11.
13. Lynch HT, Shaw MW, Magnuson CW, Larsen AL, Krush AJ. Hereditary factors in cancer. Study of two large midwestern kindreds. Arch Intern Med 1966;117:206-12.
14. Poissonnier M, Andrieu J, Gardon JD, Saint-Martin JL, Gruyer P, Martignon C. [Familial cancer of the colon without polyposis and the familial cancer syndrome. Apropos of 2 cases over 3 generations]. J Genet Hum 1983;31:255-78.
15. Bale SJ, Chakravarti A, Strong LC. Aggregation of colon cancer in family data. Genet Epidemiol 1984;1:53-61.
16. Boland CR, Troncale FJ. Familial colonic cancer without antecedent polyposis. Ann Intern Med 1984;100:700-1.
17. Rijcken FE, Hollema H, Kleibeuker JH. Proximal adenomas in hereditary non-polyposis colorectal cancer are prone to rapid malignant transformation. Gut 2002;50:382-6.
18. Umar A, Boyer JC, Thomas DC, Nguyen DC, Risinger JI, Boyd J, Ionov Y, Perucho M, Kunkel TA. Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. J Biol Chem 1994;269:14367-70.
19. Lynch HT, Smyrk T, Lynch J. An update of HNPCC (Lynch syndrome). Cancer Genet Cytogenet 1997;93:84-99.
20. Aarnio M, Salovaara R, Aaltonen LA, Mecklin JP, Jarvinen HJ. Features of gastric cancer in hereditary non-polyposis colorectal cancer syndrome. Int J Cancer 1997;74:551-5.
21. Ichikawa Y, Lemon SJ, Wang S, Franklin B, Watson P, Knezetic JA, Bewtra C, Lynch HT. Microsatellite instability and expression of MLH1 and MSH2 in normal and malignant endometrial and ovarian epithelium in hereditary nonpolyposis colorectal cancer family members. Cancer Genet Cytogenet 1999;112:2-8.
22. Green SE, Bradburn DM, Varma JS, Burn J. Hereditary non-polyposis colorectal cancer. Int J Colorectal Dis 1998;13:3-12.
23. Watson P, Vasen HF, Mecklin JP, Jarvinen H, Lynch HT. The risk of endometrial cancer in hereditary nonpolyposis colorectal cancer. Am J Med 1994;96:516-20.
24. Watson P, Lynch HT. Extracolonic cancer in hereditary nonpolyposis colorectal cancer. Cancer 1993;71:677-85.
25. Walsh JM, Terdiman JP. Colorectal cancer screening: scientific review. Jama 2003;289:1288-96.
26. Burke W, Petersen G, Lynch P, Botkin J, Daly M, Garber J, Kahn MJ, McTiernan A, Offit K, Thomson E, Varricchio C. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. I. Hereditary nonpolyposis colon cancer. Cancer Genetics Studies Consortium. Jama 1997;277:915-9.
27. Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, Mulrow CD, Woolf SH, Glick SN, Ganiats TG, Bond JH, Rosen L, Zapka JG, Olsen SJ, Giardiello FM, Sisk JE, Van Antwerp R, Brown-Davis C, Marciniak DA, Mayer RJ. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 1997;112:594-642.
28. Kruse R, Rutten A, Lamberti C, Hosseiny-Malayeri HR, Wang Y, Ruelfs C, Jungck M, Mathiak M, Ruzicka T, Hartschuh W, Bisceglia M, Friedl W, Propping P. Muir-Torre phenotype has a frequency of DNA

mismatch-repair-gene mutations similar to that in hereditary nonpolyposis colorectal cancer families defined by the Amsterdam criteria. Am J Hum Genet 1998;63:63-70.
29. Vasen HF, Sanders EA, Taal BG, Nagengast FM, Griffioen G, Menko FH, Kleibeuker JH, Houwing-Duistermaat JJ, Meera Khan P. The risk of brain tumours in hereditary non-polyposis colorectal cancer (HNPCC). Int J Cancer 1996;65:422-5.
30. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Dis Colon Rectum 1991;34:424-5.
31. Rodriguez-Bigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan PM, Lynch H, Perucho M, Smyrk T, Sobin L, Srivastava S. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst 1997;89:1758-62.
32. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 1999;116:1453-6.
33. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, 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.
34. Modrich P. Mechanisms and biological effects of mismatch repair. Annu Rev Genet 1991;25:229-53.
35. Su SS, Modrich P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. Proc Natl Acad Sci U S A 1986;83:5057-61.
36. Lahue RS, Au KG, Modrich P. DNA mismatch correction in a defined system. Science 1989;245:160-4.
37. Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994;263:1625-9.
38. Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ, Jiricny J. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science 1995;268:1912-4.
39. 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 1994;77:167.
40. Acharya S, Wilson T, Gradia S, Kane MF, Guerrette S, Marsischky GT, Kolodner R, Fishel R. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proc Natl Acad Sci U S A 1996;93:13629-34.
41. Drummond JT, Li GM, Longley MJ, Modrich P. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science 1995;268:1909-12.
42. Genschel J, Littman SJ, Drummond JT, Modrich P. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. J Biol Chem 1998;273:19895-901.
43. Palombo F, Iaccarino I, Nakajima E, Ikejima M, Shimada T, Jiricny J. hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. Curr Biol 1996;6:1181-4.
44. Buermeyer AB, Deschenes SM, Baker SM, Liskay RM. Mammalian DNA mismatch repair. Annu Rev Genet 1999;33:533-64.
45. Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A, Griffith J, Fishel R. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. Mol Cell 1999;3:255-61.
46. Fishel R. Signaling mismatch repair in cancer. Nat Med 1999;5:1239-41.
47. Sancar A. DNA repair in humans. Annu Rev Genet 1995;29:69-105.
48. Genschel J, Bazemore LR, Modrich P. Human exonuclease I is required for 5' and 3' mismatch repair. J Biol Chem 2002;277:13302-11.
49. Tomkinson AE, Mackey ZB. Structure and function of mammalian DNA ligases. Mutat Res 1998;407:1-9.
50. Drummond JT, Genschel J, Wolf E, Modrich P. DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSalpha/hMutSbeta ratio and reduces the efficiency of base-base mismatch repair. Proc Natl Acad Sci U S A 1997;94:10144-9.
51. Risinger JI, Umar A, Boyd J, Berchuck A, Kunkel TA, Barrett JC. Mutation of MSH3 in endometrial cancer and evidence for its functional role in heteroduplex repair. Nat Genet 1996;14:102-5.
52. Marra G, Iaccarino I, Lettieri T, Roscilli G, Delmastro P, Jiricny J. Mismatch repair deficiency associated with overexpression of the MSH3 gene. Proc Natl Acad Sci U S A 1998;95:8568-73.
53. Chang DK, Ricciardiello L, Goel A, Chang CL, Boland CR. Steady-state regulation of the human DNA mismatch repair system. J Biol Chem 2000;275:29178.
54. Plaschke J, Kruger S, Pistorius S, Theissig F, Saeger HD, Schackert HK. Involvement of hMSH6 in the development of hereditary and sporadic colorectal cancer revealed by immunostaining is based on germline mutations, but rarely on somatic inactivation. Int J Cancer 2002;97:643-8.

55. Montera M, Resta N, Simone C, Guanti G, Marchese C, Civitelli S, Mancini A, Pozzi S, De Salvo L, Bruzzone D, Donadini A, Romio L, Mareni C. Mutational germline analysis of hMSH2 and hMLH1 genes in early onset colorectal cancer patients. J Med Genet 2000;37:E7.
56. Bocker T, Barusevicius A, Snowden T, Rasio D, Guerrette S, Robbins D, Schmidt C, Burczak J, Croce CM, Copeland T, Kovatich AJ, Fishel R. hMSH5: a human MutS homologue that forms a novel heterodimer with hMSH4 and is expressed during spermatogenesis. Cancer Res 1999;59:816-22.
57. Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Res 2002;62:3925-8.
58. Miyakura Y, Sugano K, Akasu T, Yoshida T, Maekawa M, Saitoh S, Sasaki H, Nomizu T, Konishi F, Fujita S, Moriya Y, Nagai H. Extensive but hemiallelic methylation of the hMLH1 promoter region in early-onset sporadic colon cancers with microsatellite instability. Clin Gastroenterol Hepatol 2004;2:147-56.
59. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004;36:497-501.
60. Hitchins M, Williams R, Cheong K, Halani N, Lin VA, Packham D, Ku S, Buckle A, Hawkins N, Burn J, Gallinger S, Goldblatt J, Kirk J, Tomlinson I, Scott R, Spigelman A, Suter C, Martin D, Suthers G, Ward R. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;129:1392-9.
61. Chan TL, Yuen ST, Kong CK, Chan YW, Chan AS, Ng WF, Tsui WY, Lo MW, Tam WY, Li VS, Leung SY. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet 2006.
62. Ito E, Yanagisawa Y, Iwahashi Y, Suzuki Y, Nagasaki H, Akiyama Y, Sugano S, Yuasa Y, Maruyama K. A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene hMLH1. Biochem Biophys Res Commun 1999;256:488-94.
63. Shin YK, Heo SC, Shin JH, Hong SH, Ku JL, Yoo BC, Kim IJ, Park JG. Germline mutations in MLH1, MSH2 and MSH6 in Korean hereditary non-polyposis colorectal cancer families. Hum Mutat 2004;24:351.
64. Gazzoli I, Kolodner RD. Regulation of the human MSH6 gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms. Mol Cell Biol 2003;23:7992-8007.
65. Chung DC, Rustgi AK. DNA mismatch repair and cancer. Gastroenterology 1995;109:1685-99.
66. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, Fan RS, Zborowska E, Kinzler KW, Vogelstein B, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science 1995;268:1336-8.
67. Laken SJ, Petersen GM, Gruber SB, Oddoux C, Ostrer H, Giardiello FM, Hamilton SR, Hampel H, Markowitz A, Klimstra D, Jhanwar S, Winawer S, Offit K, Luce MC, Kinzler KW, Vogelstein B. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. Nat Genet 1997;17:79-83.
68. 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;275:967-9.
69. Malkhosyan S, Rampino N, Yamamoto H, Perucho M. Frameshift mutator mutations. Nature 1996;382:499-500.
70. Yamamoto H, Sawai H, Perucho M. Frameshift somatic mutations in gastrointestinal cancer of the microsatellite mutator phenotype. Cancer Res 1997;57:4420-6.
71. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J, Miliaresis C, Rodgers L, McCombie R, Bigner SH, Giovanella BC, Ittmann M, Tycko B, Hibshoosh H, Wigler MH, Parsons R. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 1997;275:1943-7.
72. Souza RF, Yin J, Smolinski KN, Zou TT, Wang S, Shi YQ, Rhyu MG, Cottrell J, Abraham JM, Biden K, Simms L, Leggett B, Bova GS, Frank T, Powell SM, Sugimura H, Young J, Harpaz N, Shimizu K, Matsubara N, Meltzer SJ. Frequent mutation of the E2F-4 cell cycle gene in primary human gastrointestinal tumors. Cancer Res 1997;57:2350-3.
73. 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;363:558-61.
74. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993;260:816-9.
75. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Powell SM, Jen J, Hamilton SR, et al. Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812-6.
76. Moslein G, Tester DJ, Lindor NM, Honchel R, Cunningham JM, French AJ, Halling KC, Schwab M, Goretzki P, Thibodeau SN. Microsatellite instability and mutation analysis of hMSH2 and hMLH1 in patients with sporadic, familial and hereditary colorectal cancer. Hum Mol Genet 1996;5:1245-52.
77. Liu B, Nicolaides NC, Markowitz S, Willson JK, Parsons RE, Jen J, Papadopolous N, Peltomaki P, de la Chapelle A, Hamilton SR, et al. Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. Nat Genet 1995;9:48-55.

78. Miozzo M, Sozzi G, Musso K, Pilotti S, Incarbone M, Pastorino U, Pierotti MA. Microsatellite alterations in bronchial and sputum specimens of lung cancer patients. Cancer Res 1996;56:2285-8.
79. Karnik P, Plummer S, Casey G, Myles J, Tubbs R, Crowe J, Williams BR. Microsatellite instability at a single locus (D11S988) on chromosome 11p15.5 as a late event in mammary tumorigenesis. Hum Mol Genet 1995;4:1889-94.
80. dos Santos NR, Seruca R, Constancia M, Seixas M, Sobrinho-Simoes M. Microsatellite instability at multiple loci in gastric carcinoma: clinicopathologic implications and prognosis. Gastroenterology 1996;110:38-44.
81. Han HJ, Yanagisawa A, Kato Y, Park JG, Nakamura Y. Genetic instability in pancreatic cancer and poorly differentiated type of gastric cancer. Cancer Res 1993;53:5087-9.
82. Risinger JI, Umar A, Boyer JC, Evans AC, Berchuck A, Kunkel TA, Barrett JC. Microsatellite instability in gynecological sarcomas and in hMSH2 mutant uterine sarcoma cell lines defective in mismatch repair activity. Cancer Res 1995;55:5664-9.
83. Kobayashi K, Matsushima M, Koi S, Saito H, Sagae S, Kudo R, Nakamura Y. Mutational analysis of mismatch repair genes, hMLH1 and hMSH2, in sporadic endometrial carcinomas with microsatellite instability. Jpn J Cancer Res 1996;87:141-5.
84. Fujita M, Enomoto T, Yoshino K, Nomura T, Buzard GS, Inoue M, Okudaira Y. Microsatellite instability and alterations in the hMSH2 gene in human ovarian cancer. Int J Cancer 1995;64:361-6.
85. Aaltonen LA, Peltomaki P, Mecklin JP, Jarvinen H, Jass JR, Green JS, Lynch HT, Watson P, Tallqvist G, Juhola M, et al. Replication errors in benign and malignant tumors from hereditary nonpolyposis colorectal cancer patients. Cancer Res 1994;54:1645-8.
86. Ward R, Meagher A, Tomlinson I, O'Connor T, Norrie M, Wu R, Hawkins N. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. Gut 2001;48:821-9.
87. 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.
88. Umar A, Koi M, Risinger JI, Glaab WE, Tindall KR, Kolodner RD, Boland CR, Barrett JC, Kunkel TA. Correction of hypermutability, N-methyl-N'-nitro-N-nitrosoguanidine resistance, and defective DNA mismatch repair by introducing chromosome 2 into human tumor cells with mutations in MSH2 and MSH6. Cancer Res 1997;57:3949-55.
89. Andrew SE, Reitmair AH, Fox J, Hsiao L, Francis A, McKinnon M, Mak TW, Jirik FR. Base transitions dominate the mutational spectrum of a transgenic reporter gene in MSH2 deficient mice. Oncogene 1997;15:123-9.
90. Toft NJ, Winton DJ, Kelly J, Howard LA, Dekker M, te Riele H, Arends MJ, Wyllie AH, Margison GP, Clarke AR. Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. Proc Natl Acad Sci U S A 1999;96:3911-5.
91. Alani E. The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. Mol Cell Biol 1996;16:5604-15.
92. Marsischky GT, Filosi N, Kane MF, Kolodner R. Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev 1996;10:407-20.
93. Lynch HT, Lynch JF. Lynch syndrome: history and current status. Dis Markers 2004;20:181-98.
94. Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Peltomaki P, de la Chapelle A, Hamilton SR, Vogelstein B, Kinzler KW. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat Med 1996;2:169-74.
95. Nystrom-Lahti M, Wu Y, Moisio AL, Hofstra RM, Osinga J, Mecklin JP, Jarvinen HJ, Leisti J, Buys CH, de la Chapelle A, Peltomaki P. DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum Mol Genet 1996;5:763-9.
96. Aaltonen LA, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomaki P, Chadwick RB, Kaariainen H, Eskelinen M, Jarvinen H, Mecklin JP, de la Chapelle A. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 1998;338:1481-7.
97. Peltomaki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. J Clin Oncol 2003;21:1174-9.
98. Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 2004;20:269-76.
99. Peltomaki P, Vasen HF. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology 1997;113:1146-58.
100. Charbonnier F, Raux G, Wang Q, Drouot N, Cordier F, Limacher JM, Saurin JC, Puisieux A, Olschwang S, Frebourg T. Detection of exon deletions and duplications of the mismatch repair genes in hereditary

nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. Cancer Res 2000;60:2760-3.
101. Charbonnier F, Olschwang S, Wang Q, Boisson C, Martin C, Buisine MP, Puisieux A, Frebourg T. MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. Cancer Res 2002;62:848-53.
102. Gille JJ, Hogervorst FB, Pals G, Wijnen JT, van Schooten RJ, Dommering CJ, Meijer GA, Craanen ME, Nederlof PM, de Jong D, McElgunn CJ, Schouten JP, Menko FH. Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. Br J Cancer 2002;87:892-7.
103. Di Fiore F, Charbonnier F, Martin C, Frerot S, Olschwang S, Wang Q, Boisson C, Buisine MP, Nilbert M, Lindblom A, Frebourg T. Screening for genomic rearrangements of the MMR genes must be included in the routine diagnosis of HNPCC. J Med Genet 2004;41:18-20.
104. van der Klift H, Wijnen J, Wagner A, Verkuilen P, Tops C, Otway R, Kohonen-Corish M, Vasen H, Oliani C, Barana D, Moller P, Delozier-Blanchet C, Hutter P, Foulkes W, Lynch H, Burn J, Moslein G, Fodde R. Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC). Genes Chromosomes Cancer 2005.
105. Moisio AL, Sistonen P, Weissenbach J, de la Chapelle A, Peltomaki P. Age and origin of two common MLH1 mutations predisposing to hereditary colon cancer. Am J Hum Genet 1996;59:1243-51.
106. Jager AC, Bisgaard ML, Myrhoj T, Bernstein I, Rehfeld JF, Nielsen FC. Reduced frequency of extracolonic cancers in hereditary nonpolyposis colorectal cancer families with monoallelic hMLH1 expression. Am J Hum Genet 1997;61:129-38.
107. Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. J Med Genet 1999;36:801-18.
108. Foulkes WD, Thiffault I, Gruber SB, Horwitz M, Hamel N, Lee C, Shia J, Markowitz A, Figer A, Friedman E, Farber D, Greenwood CM, Bonner JD, Nafa K, Walsh T, Marcus V, Tomsho L, Gebert J, Macrae FA, Gaff CL, Paillerets BB, Gregersen PK, Weitzel JN, Gordon PH, MacNamara E, King MC, Hampel H, De La Chapelle A, Boyd J, Offit K, Rennert G, Chong G, Ellis NA. The founder mutation MSH2*1906G-->C is an important cause of hereditary nonpolyposis colorectal cancer in the Ashkenazi Jewish population. Am J Hum Genet 2002;71:1395-412.
109. Guillem JG, Rapaport BS, Kirchhoff T, Kolachana P, Nafa K, Glogowski E, Finch R, Huang H, Foulkes WD, Markowitz A, Ellis NA, Offit K. A636P is associated with early-onset colon cancer in Ashkenazi Jews. J Am Coll Surg 2003;196:222-5.
110. Froggatt NJ, Green J, Brassett C, Evans DG, Bishop DT, Kolodner R, Maher ER. A common MSH2 mutation in English and North American HNPCC families: origin, phenotypic expression, and sex specific differences in colorectal cancer. J Med Genet 1999;36:97-102.
111. Wagner A, Barrows A, Wijnen JT, van der Klift H, Franken PF, Verkuijlen P, Nakagawa H, Geugien M, Jaghmohan-Changur S, Breukel C, Meijers-Heijboer H, Morreau H, van Puijenbroek M, Burn J, Coronel S, Kinarski Y, Okimoto R, Watson P, Lynch JF, de la Chapelle A, Lynch HT, Fodde R. Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the MSH2 gene. Am J Hum Genet 2003;72:1088-100.
112. Lynch HT, Coronel SM, Okimoto R, Hampel H, Sweet K, Lynch JF, Barrows A, Wijnen J, van der Klift H, Franken P, Wagner A, Fodde R, de la Chapelle A. A founder mutation of the MSH2 gene and hereditary nonpolyposis colorectal cancer in the United States. Jama 2004;291:718-24.
113. Miyaki M, Konishi M, Tanaka K, Kikuchi-Yanoshita R, Muraoka M, Yasuno M, Igari T, Koike M, Chiba M, Mori T. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet 1997;17:271-2.
114. Akiyama Y, Sato H, Yamada T, Nagasaki H, Tsuchiya A, Abe R, Yuasa Y. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. Cancer Res 1997;57:3920-3.
115. Wu Y, Berends MJ, Mensink RG, Kempinga C, Sijmons RH, van Der Zee AG, Hollema H, Kleibeuker JH, Buys CH, Hofstra RM. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. Am J Hum Genet 1999;65:1291-8.
116. Wijnen J, de Leeuw W, Vasen H, van der Klift H, Moller P, Stormorken A, Meijers-Heijboer H, Lindhout D, Menko F, Vossen S, Moslein G, Tops C, Brocker-Vriends A, Wu Y, Hofstra R, Sijmons R, Cornelisse C, Morreau H, Fodde R. Familial endometrial cancer in female carriers of MSH6 germline mutations. Nat Genet 1999;23:142-4.
117. Wagner A, Hendriks Y, Meijers-Heijboer EJ, de Leeuw WJ, Morreau H, Hofstra R, Tops C, Bik E, Brocker-Vriends AH, van Der Meer C, Lindhout D, Vasen HF, Breuning MH, Cornelisse CJ, van Krimpen C,

Niermeijer MF, Zwinderman AH, Wijnen J, Fodde R. Atypical HNPCC owing to MSH6 germline mutations: analysis of a large Dutch pedigree. J Med Genet 2001;38:318-22.
118. Kolodner RD, Tytell JD, Schmeits JL, Kane MF, Gupta RD, Weger J, Wahlberg S, Fox EA, Peel D, Ziogas A, Garber JE, Syngal S, Anton-Culver H, Li FP. Germ-line msh6 mutations in colorectal cancer families. Cancer Res 1999;59:5068-74.
119. Berends MJ, Wu Y, Sijmons RH, Mensink RG, van der Sluis T, Hordijk-Hos JM, de Vries EG, Hollema H, Karrenbeld A, Buys CH, van der Zee AG, Hofstra RM, Kleibeuker JH. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. Am J Hum Genet 2002;70:26-37.
120. Huang J, Kuismanen SA, Liu T, Chadwick RB, Johnson CK, Stevens MW, Richards SK, Meek JE, Gao X, Wright FA, Mecklin JP, Jarvinen HJ, Gronberg H, Bisgaard ML, Lindblom A, Peltomaki P. MSH6 and MSH3 are rarely involved in genetic predisposition to nonpolypotic colon cancer. Cancer Res 2001;61:1619-23.
121. Papadopoulos N, Nicolaides NC, Liu B, Parsons R, Lengauer C, Palombo F, D'Arrigo A, Markowitz S, Willson JK, Kinzler KW, et al. Mutations of GTBP in genetically unstable cells. Science 1995;268:1915-7.
122. Hendriks YM, Wagner A, Morreau H, Menko F, Stormorken A, Quehenberger F, Sandkuijl L, Moller P, Genuardi M, Van Houwelingen H, Tops C, Van Puijenbroek M, Verkuijlen P, Kenter G, Van Mil A, Meijers-Heijboer H, Tan GB, Breuning MH, Fodde R, Wijnen JT, Brocker-Vriends AH, Vasen H. Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counseling and surveillance. Gastroenterology 2004;127:17-25.
123. de Leeuw WJ, Dierssen J, Vasen HF, Wijnen JT, Kenter GG, Meijers-Heijboer H, Brocker-Vriends A, Stormorken A, Moller P, Menko F, Cornelisse CJ, Morreau H. Prediction of a mismatch repair gene defect by microsatellite instability and immunohistochemical analysis in endometrial tumours from HNPCC patients. J Pathol 2000;192:328-35.
124. Edelmann W, Yang K, Umar A, Heyer J, Lau K, Fan K, Liedtke W, Cohen PE, Kane MF, Lipford JR, Yu N, Crouse GF, Pollard JW, Kunkel T, Lipkin M, Kolodner R, Kucherlapati R. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell 1997;91:467-77.
125. de Wind N, Dekker M, Claij N, Jansen L, van Klink Y, Radman M, Riggins G, van der Valk M, van't Wout K, te Riele H. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet 1999;23:359-62.
126. Plaschke J, Ruschoff J, Schackert HK. Genomic rearrangements of hMSH6 contribute to the genetic predisposition in suspected hereditary non-polyposis colorectal cancer syndrome. J Med Genet 2003;40:597-600.
127. Parc YR, Halling KC, Wang L, Christensen ER, Cunningham JM, French AJ, Burgart LJ, Price-Troska TL, Roche PC, Thibodeau SN. HMSH6 alterations in patients with microsatellite instability-low colorectal cancer. Cancer Res 2000;60:2225-31.
128. Verma L, Kane MF, Brassett C, Schmeits J, Evans DG, Kolodner RD, Maher ER. Mononucleotide microsatellite instability and germline MSH6 mutation analysis in early onset colorectal cancer. J Med Genet 1999;36:678-82.
129. Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75-80.
130. Liu T, Yan H, Kuismanen S, Percesepe A, Bisgaard ML, Pedroni M, Benatti P, Kinzler KW, Vogelstein B, Ponz de Leon M, Peltomaki P, Lindblom A. The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. Cancer Res 2001;61:7798-802.
131. Wu Y, Berends MJ, Sijmons RH, Mensink RG, Verlind E, Kooi KA, van der Sluis T, Kempinga C, van dDer Zee AG, Hollema H, Buys CH, Kleibeuker JH, Hofstra RM. A role for MLH3 in hereditary nonpolyposis colorectal cancer. Nat Genet 2001;29:137-8.
132. Hienonen T, Laiho P, Salovaara R, Mecklin JP, Jarvinen H, Sistonen P, Peltomaki P, Lehtonen R, Nupponen NN, Launonen V, Karhu A, Aaltonen LA. Little evidence for involvement of MLH3 in colorectal cancer predisposition. Int J Cancer 2003;106:292-6.
133. Liu HX, Zhou XL, Liu T, Werelius B, Lindmark G, Dahl N, Lindblom A. The role of hMLH3 in familial colorectal cancer. Cancer Res 2003;63:1894-9.
134. Amin NS, Nguyen MN, Oh S, Kolodner RD. exo1-Dependent mutator mutations: model system for studying functional interactions in mismatch repair. Mol Cell Biol 2001;21:5142-55.
135. Nakagawa H, Lockman JC, Frankel WL, Hampel H, Steenblock K, Burgart LJ, Thibodeau SN, de la Chapelle A. Mismatch repair gene PMS2: disease-causing germline mutations are frequent in patients whose tumors stain negative for PMS2 protein, but paralogous genes obscure mutation detection and interpretation. Cancer Res 2004;64:4721-7.
136. Truninger K, Menigatti M, Luz J, Russell A, Haider R, Gebbers JO, Bannwart F, Yurtsever H, Neuweiler J, Riehle HM, Cattaruzza MS, Heinimann K, Schar P, Jiricny J, Marra G. Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer. Gastroenterology 2005;128:1160-71.

137. Worthley DL, Walsh MD, Barker M, Ruszkiewicz A, Bennett G, Phillips K, Suthers G. Familial mutations in PMS2 can cause autosomal dominant hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;128:1431-6.
138. De Rosa M, Fasano C, Panariello L, Scarano MI, Belli G, Iannelli A, Ciciliano F, Izzo P. Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. Oncogene 2000;19:1719-23.
139. De Vos M, Hayward BE, Picton S, Sheridan E, Bonthron DT. Novel PMS2 pseudogenes can conceal recessive mutations causing a distinctive childhood cancer syndrome. Am J Hum Genet 2004;74:954-64.
140. Hendriks YM, Jagmohan-Changur S, van der Klift HM, Morreau H, van Puijenbroek M, Tops C, van Os T, Wagner A, Ausems MG, Gomez E, Breuning MH, Brocker-Vriends AH, Vasen HF, Wijnen JT. Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). Gastroenterology 2006;130:312-22.
141. Herman JG, Civin CI, Issa JP, Collector MI, Sharkis SJ, Baylin SB. Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies. Cancer Res 1997;57:837-41.
142. Baylin SB, Makos M, Wu JJ, Yen RW, de Bustros A, Vertino P, Nelkin BD. Abnormal patterns of DNA methylation in human neoplasia: potential consequences for tumor progression. Cancer Cells 1991;3:383-90.
143. Gonzalgo ML, Hayashida T, Bender CM, Pao MM, Tsai YC, Gonzales FA, Nguyen HD, Nguyen TT, Jones PA. The role of DNA methylation in expression of the p19/p16 locus in human bladder cancer cell lines. Cancer Res 1998;58:1245-52.
144. 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.
145. Kuismanen SA, Holmberg MT, Salovaara R, de la Chapelle A, 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.
146. Furukawa T, Konishi F, Masubuchi S, Shitoh K, Nagai H, Tsukamoto T. Densely methylated MLH1 promoter correlates with decreased mRNA expression in sporadic colorectal cancers. Genes Chromosomes Cancer 2002;35:1-10.
147. Menigatti M, Di Gregorio C, Borghi F, Sala E, Scarselli A, Pedroni M, Foroni M, Benatti P, Roncucci L, Ponz de Leon M, Percesepe A. Methylation pattern of different regions of the MLH1 promoter and silencing of gene expression in hereditary and sporadic colorectal cancer. Genes Chromosomes Cancer 2001;31:357-61.
148. Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD, Markowitz SD. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95:8698-702.
149. 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.
150. Esteller M, Levine R, Baylin SB, Ellenson LH, Herman JG. MLH1 promoter hypermethylation is associated with the microsatellite instability phenotype in sporadic endometrial carcinomas. Oncogene 1998;17:2413-7.
151. Esteller M, Catasus L, Matias-Guiu X, Mutter GL, Prat J, Baylin SB, Herman JG. hMLH1 promoter hypermethylation is an early event in human endometrial tumorigenesis. Am J Pathol 1999;155:1767-72.
152. Simpkins SB, Bocker T, Swisher EM, Mutch DG, Gersell DJ, Kovatich AJ, Palazzo JP, Fishel R, Goodfellow PJ. MLH1 promoter methylation and gene silencing is the primary cause of microsatellite instability in sporadic endometrial cancers. Hum Mol Genet 1999;8:661-6.
153. Berends MJ, Hollema H, Wu Y, van Der Sluis T, Mensink RG, ten Hoor KA, Sijmons RH, de Vries EG, Pras E, Mourits MJ, Hofstra RM, Buys CH, Kleibeuker JH, van Der Zee AG. MLH1 and MSH2 protein expression as a pre-screening marker in hereditary and non-hereditary endometrial hyperplasia and cancer. Int J Cancer 2001;92:398-403.
154. Baldinu P, Cossu A, Manca A, Satta MP, Pisano M, Casula M, Dessole S, Pintus A, Tanda F, Palmieri G. Microsatellite instability and mutation analysis of candidate genes in unselected sardinian patients with endometrial carcinoma. Cancer 2002;94:3157-68.
155. Eads CA, Lord RV, Wickramasinghe K, Long TI, Kurumboor SK, Bernstein L, Peters JH, DeMeester SR, DeMeester TR, Skinner KA, Laird PW. Epigenetic patterns in the progression of esophageal adenocarcinoma. Cancer Res 2001;61:3410-8.
156. Kang GH, Lee HJ, Hwang KS, Lee S, Kim JH, Kim JS. Aberrant CpG island hypermethylation of chronic gastritis, in relation to aging, gender, intestinal metaplasia, and chronic inflammation. Am J Pathol 2003;163:1551-6.

157. Jeronimo C, Usadel H, Henrique R, Oliveira J, Lopes C, Nelson WG, Sidransky D. Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. J Natl Cancer Inst 2001;93:1747-52.
158. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 2001;61:3573-7.
159. Sun Y, Deng D, You WC, Bai H, Zhang L, Zhou J, Shen L, Ma JL, Xie YQ, Li JY. Methylation of p16 CpG islands associated with malignant transformation of gastric dysplasia in a population-based study. Clin Cancer Res 2004;10:5087-93.
160. Anttila S, Hakkola J, Tuominen P, Elovaara E, Husgafvel-Pursiainen K, Karjalainen A, Hirvonen A, Nurminen T. Methylation of cytochrome P4501A1 promoter in the lung is associated with tobacco smoking. Cancer Res 2003;63:8623-8.
161. Menzelio D. De excrescentiis verrucoso cristosis copiose in intestinis crassis dysenteriam passi observatis. Acta Medicorum Berolinensium 1721;IX:68-71.
162. Cripps W. Two cases of disseminated polypus of the rectum. Trans Path Soc London 1882;33:165-8.
163. Smith T. Three cases of multiple polypi of the lower bowel occuring in one family. St. Bartholomews Hosp Rep 1887:225-9.
164. Bickersteth R. Multiple polypi of the rectum occurring in a mother and child. St. Bartholomews Hosp Rep 1890:299-301.
165. Cruz-Correa M, Giardiello FM. Diagnosis and management of hereditary colon cancer. Gastroenterol Clin North Am 2002;31:537-49, x.
166. Bulow S. Familial polyposis coli. Dan Med Bull 1987;34:1-15.
167. Jarvinen HJ. Time and type of prophylactic surgery for familial adenomatosis coli. Ann Surg 1985;202:93-7.
168. Offerhaus GJ, Giardiello FM, Krush AJ, Booker SV, Tersmette AC, Kelley NC, Hamilton SR. The risk of upper gastrointestinal cancer in familial adenomatous polyposis. Gastroenterology 1992;102:1980-2.
169. Gardner EJ. A genetic and clinical study of intestinal polyposis, a predisposing factor for carcinoma of the colon and rectum. Am J Hum Genet 1951;3:167-76.
170. Gardner EJ. Follow-up study of a family group exhibiting dominant inheritance for a syndrome including intestinal polyps, osteomas, fibromas and epidermal cysts. Am J Hum Genet 1962;14:376-90.
171. Smith WG. Multiple polyposis, Gardner's syndrome and desmoid tumors. Dis Colon Rectum 1958;1:323-32.
172. Leppard BJ. Epidermoid cysts and polyposis coli. Proc R Soc Med 1974;67:1036-7.
173. Traboulsi EI, Murphy SF, de la Cruz ZC, Maumenee IH, Green WR. A clinicopathologic study of the eyes in familial adenomatous polyposis with extracolonic manifestations (Gardner's syndrome). Am J Ophthalmol 1990;110:550-61.
174. Paraf F, Jothy S, Van Meir EG. Brain tumor-polyposis syndrome: two genetic diseases? J Clin Oncol 1997;15:2744-58.
175. Hamilton SR, Liu B, Parsons RE, Papadopoulos N, Jen J, Powell SM, Krush AJ, Berk T, Cohen Z, Tetu B, et al. The molecular basis of Turcot's syndrome. N Engl J Med 1995;332:839-47.
176. Qualman SJ, Bowen J, Erdman SH. Molecular basis of the brain tumor-polyposis (Turcot) syndrome. Pediatr Dev Pathol 2003;6:574-6.
177. Lucci-Cordisco E, Zito I, Gensini F, Genuardi M. Hereditary nonpolyposis colorectal cancer and related conditions. Am J Med Genet A 2003;122:325-34.
178. Lynch HT, Smyrk TC, Watson P, Lanspa SJ, Lynch PM, Jenkins JX, Rouse J, Cavalieri J, Howard L, Lynch J. Hereditary flat adenoma syndrome: a variant of familial adenomatous polyposis? Dis Colon Rectum 1992;35:411-21.
179. Bodmer WF, Bailey CJ, Bodmer J, Bussey HJ, Ellis A, Gorman P, Lucibello FC, Murday VA, Rider SH, Scambler P, et al. Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 1987;328:614-6.
180. Leppert M, Dobbs M, Scambler P, O'Connell P, Nakamura Y, Stauffer D, Woodward S, Burt R, Hughes J, Gardner E, et al. The gene for familial polyposis coli maps to the long arm of chromosome 5. Science 1987;238:1411-3.
181. Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hedge P, McKechnie D, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661-5.
182. Groden J, Thliveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, Joslyn G, Stevens J, Spirio L, Robertson M, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 1991;66:589-600.
183. Laurent-Puig P, Beroud C, Soussi T. APC gene: database of germline and somatic mutations in human tumors and cell lines. Nucleic Acids Res 1998;26:269-70.
184. Grady WM. Genetic testing for high-risk colon cancer patients. Gastroenterology 2003;124:1574-94.

185. Wallis YL, Morton DG, McKeown CM, Macdonald F. Molecular analysis of the APC gene in 205 families: extended genotype-phenotype correlations in FAP and evidence for the role of APC amino acid changes in colorectal cancer predisposition. J Med Genet 1999;36:14-20.
186. Spirio L, Olschwang S, Groden J, Robertson M, Samowitz W, Joslyn G, Gelbert L, Thliveris A, Carlson M, Otterud B, et al. Alleles of the APC gene: an attenuated form of familial polyposis. Cell 1993;75:951-7.
187. Giardiello FM, Brensinger JD, Luce MC, Petersen GM, Cayouette MC, Krush AJ, Bacon JA, Booker SV, Bufill JA, Hamilton SR. Phenotypic expression of disease in families that have mutations in the 5' region of the adenomatous polyposis coli gene. Ann Intern Med 1997;126:514-9.
188. Brensinger JD, Laken SJ, Luce MC, Powell SM, Vance GH, Ahnen DJ, Petersen GM, Hamilton SR, Giardiello FM. Variable phenotype of familial adenomatous polyposis in pedigrees with 3' mutation in the APC gene. Gut 1998;43:548-52.
189. Nagase H, Miyoshi Y, Horii A, Aoki T, Ogawa M, Utsunomiya J, Baba S, Sasazuki T, Nakamura Y. Correlation between the location of germ-line mutations in the APC gene and the number of colorectal polyps in familial adenomatous polyposis patients. Cancer Res 1992;52:4055-7.
190. Nugent KP, Phillips RK, Hodgson SV, Cottrell S, Smith-Ravin J, Pack K, Bodmer WF. Phenotypic expression in familial adenomatous polyposis: partial prediction by mutation analysis. Gut 1994;35:1622-3.
191. Lamlum H, Ilyas M, Rowan A, Clark S, Johnson V, Bell J, Frayling I, Efstathiou J, Pack K, Payne S, Roylance R, Gorman P, Sheer D, Neale K, Phillips R, Talbot I, Bodmer W, Tomlinson I. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. Nat Med 1999;5:1071-5.
192. Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, Papadopoulou A, Bicknell D, Bodmer WF, Tomlinson IP. APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". Proc Natl Acad Sci U S A 2000;97:3352-7.
193. Albuquerque C, Breukel C, van der Luijt R, Fidalgo P, Lage P, Slors FJ, Leitao CN, Fodde R, Smits R. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. Hum Mol Genet 2002;11:1549-60.
194. Davies DR, Armstrong JG, Thakker N, Horner K, Guy SP, Clancy T, Sloan P, Blair V, Dodd C, Warnes TW, et al. Severe Gardner syndrome in families with mutations restricted to a specific region of the APC gene. Am J Hum Genet 1995;57:1151-8.
195. Olschwang S, Tiret A, Laurent-Puig P, Muleris M, Parc R, Thomas G. Restriction of ocular fundus lesions to a specific subgroup of APC mutations in adenomatous polyposis coli patients. Cell 1993;75:959-68.
196. Caspari R, Olschwang S, Friedl W, Mandl M, Boisson C, Boker T, Augustin A, Kadmon M, Moslein G, Thomas G, et al. Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions (CHRPE) associated with APC mutations beyond codon 1444. Hum Mol Genet 1995;4:337-40.
197. Woodage T, King SM, Wacholder S, Hartge P, Struewing JP, McAdams M, Laken SJ, Tucker MA, Brody LC. The APCI1307K allele and cancer risk in a community-based study of Ashkenazi Jews. Nat Genet 1998;20:62-5.
198. Gryfe R, Di Nicola N, Lal G, Gallinger S, Redston M. Inherited colorectal polyposis and cancer risk of the APC I1307K polymorphism. Am J Hum Genet 1999;64:378-84.

Chapter 2

Molecular screening of families affected by familial adenomatous polyposis (FAP)

Gazzoli I, De Andreis C, Sirchia SM, Sala P, Rossetti C, Bertario L and Colucci G.

Journal of Medical Screening 1996; 3(4):195-9.

Abstract

Objectives-To assess the risk of developing familial adenomatous polyposis (FAP) in presymptomatic individuals using APC gene flanking and intragenic polymorphic markers.

Setting-Twenty families enrolled in the Italian Registry of Polyposis comprising a total of 217 individuals, including 53 (24%) presymptomatic subjects with a 50% priori risk of FAP, were analyzed. Direct analysis techniques had previously failed to identify the FAP mutations in these families.

Methods-DNA isolated from peripheral mononuclear blood cells and tissue sections was analyzed by the polymerase chain reaction and a panel of seven highly polymorphic markers- YN5.64, CB83, CB26, LNS, APC 1458.5, MBC, 37AB. Amplification products were separated by a modified denaturing gel electrophoresis method.

Results-The haplotype associated with the disease was identified in 18 families (90%). The segregation of the FAP haplotype in these kindreds showed that 10 presymptomatic individuals had inherited the FAP mutation and carried a high risk of developing the disease. The remaining two families were not informative because of the lack of a sufficient number of probands or biological specimens.

Conclusions-These data indicate that indirect analysis with linked DNA markers has a high rate of success in defining the risk of FAP of presymptomatic subjects, provided that a sufficient number of probands or samples are available. Uninformative families accounted for 10% of the total, indicating that linkage analysis may still have higher sensitivity than direct mutation analysis techniques. The combined use of both approaches should be implemented, however, to enhance further the application of molecular genetics to the screening of families with FAP.

Introduction

Familial adenomatous polyposis (FAP) is an inherited autosomal, dominant disease characterized by the development of hundreds to thousands of polyposis in the colon and rectum. If left without surgical treatment, one or more of these benign lesions will inevitably progress to cancer [1-3]. Development of polyposis in the bowels of patients with FAP is usually associated with extracolonic manifestations, such as osteomas, desmoid, tumors, and various other neoplasms [4]. The children of patients with FAP have a 50% chance of inheriting the defective gene and developing colonic polyps. The age of onset of the disease varies, but it is recommended that all people at risk are screened by repeated sigmoidoscopy from puberty until at least the age of 40.

A few years ago the genetic defect of FAP was mapped to human chromosome 5q21-22 and in 1991 *Groden et al* cloned and sequenced the APC (adenomatous polyposis coli) gene [5-7]**.** The APC transcript contains 8535 nucleotides divided in 15 exons, the last including more than three quarters of the coding sequence [7, 8]**.**

Identification of polymorphic markers closely related to the APC gene allowed for the introduction of non-invasive techniques based on molecular diagnostic procedures that can reduce the frequency of endoscopic screening of presymptomatic individuals [9-12].

Moreover, by direct sequencing several mutations have been found within the APC gene, most of which cause the synthesis of truncated proteins detectable using the protein truncation technique (PTT) [13-18]. However, as mutations can be successfully identified in only 50-60% of cases, indirect linkage analysis of affected families is still a useful tool to identify the risk of presymptomatic individuals [19].

Owing to the poor sensitivity of mutation detection techniques and the length of time needed for the analysis, many centers offering testing for FAP use linkage analysis as first line screen. Most families are suitable as the new mutation rate is relatively low (because there are few families with only a single affected individual). Mutation
detection requires analysis of the smaller exons (2-14) individually by techniques such as single strand conformation polymorphism (SSCP), and the large exon 15, even with rapid detection

techniques such as PTT, may still require several separate reactions. Screening for mutation may therefore take several months with only limited chance of identifying the mutation. Although some laboratories work with RNA and PTT, which requires fewer tests, most find these techniques difficult. Given the fact that linkage analysis with informative intragenic and flanking markers may have greater than 99.9% accuracy, little is sacrificed by not finding a mutation. Indeed as part of the rationale for testing is to discharge those at low risk from screening programmes, and the risk of these individuals developing a sporadic bowel cancer is more than 100 times their risk of developing an FAP cancer owing to an inaccurate prediction, there does not seem to be much lost by the small difference in accuracy between the two tests.

In this study we describe the results of a screening program we started several years ago to identify the risk of FAP in patients enrolled in the Italian Registry of Familial Polyposis.

Materials and methods

We studied a total of 217 patients from 20 families referred to the Italian Registry of familial Polyposis. They had been unsuccessfully tested by mutation analysis based on SSCP followed by direct sequencing. Patients' mean age was 45 (range 12-70). In 83 patients FAP had already been diagnosed by clinical examinations. We used a panel of polymorphic markers to identify the haplotype associated with the disease gene in each family.

DNA extraction and amplification. DNA was isolated from peripheral blood or tissue sections obtained from the 217 patients. After conventional phenol/ chloroform extraction and ethanol precipitation, DNA samples were tested on 0.4% agarose gel to determine their concentration and purity and then amplified using primers specific for seven polymorphic loci linked to the APC gene[20-25] (table1). In the case of tissue sections, sequential xylene/ethanol incubations and washes were performed to remove the paraffin embedding before DNA extraction.

200 ng of DNA was added to a solution containing 0.5 µmol of each primer, 200 µmol of each deoxynucleoside triphosphate, 10 mmol Tris-HCl (pH 8.4), 50 mmol KCl, 2.0 mmol MgCl2, 2U Taq DNA polymerase in a volume of 50 µmol, and then processed for 33 polymerase chain reaction cycles of amplification, each including denaturation at 95ºC for one minute, annealing at 58ºC for 30 seconds, and extension at 70ºC for one minute. The amplified restriction fragment length

polymorphism APC 1458.5 and 37AB were digested with RsaI and SspI respectively, under conditions specified by the manufacturer (Molecular Biology resources, Milwaukee, USA). The digested samples were electrophoresed on 3% agarose gel to visualize the different length of the restriction fragments. The amplified CA-repeats were analyzed by constant denaturing gel electrophoresis, a procedure which can separate alleles that differ buy two or more bases [26]. In brief, the amplification products were mixed with ¼ vol. 0.3% xylene, 0.3 M bromophenol blue, 1:1 100% formamide/H_20, then boiled for three minutes and, finally kept on ice until loading to avoid renaturation. The gel, comprising 10% polyacrilamide (acrylamide/bis-acrylamide 19:1), 8 M urea, and 1X TAE buffer (40mM Tris acetate, 1mM EDTA), was submerged in TAE buffer at 60ºC and the samples run at 400V for four hours. After electrophoresis, gels were stained with ethidium bromide and analyzed by ultraviolet spectroscopy.

Table 1 Main characteristics of the DNA markers used for indirect analysis

Marker	*Locus*	*Type of polymorphism*	*No of alleles*	*PIC**	*HET†*	*References*
CB83	D5S122	CA-repeat	2	0.54	—	22
YN5.64	D5S82	CA-repeat	6	0.70	—	31
MBC	MCC gene	CA-repeat	4	0.50	—	32
CB26	D5S299	CA-repeat	7	0.66	—	33
LNS	D5S346	CA-repeat	13	—	0.83	34
APC1458.5	—	RFLP‡	2	—	0.48	35
37AB	3' APC	RFLP	2	—	0.52	35

* Polymorphic information content.
† Heterozygosity rate.
‡ RFLP = restriction fragment length polymorphism.

Risk evaluation. The risk of FAP was calculated according to the Bayes's theorem of prior and posterior probability, assuming a prior probability ranging form 50% to 3% depending on the age of the subject (50% for individuals aged 10 or younger to 3% for individuals aged 30 or older) [27].

Results

Table 2 shows details of the 20 families with FAP. Families 1-18 were informative and we succeeded in estimating the risk of asymptomatic individuals. Families 19-20 had unsuitable pedigree structure for indirect analysis; there were too few probands and/or biological specimens to allow completion of the analysis. Table 2 shows that 10 out 53 presymptomatic individuals had

inherited the FAP associated haplotype in addition to other clinical parameters, including CHRPE (congenital hypertrophy retinal pigmented epithelium), osteomas, and colonoscopy.

Table 2 Results of the indirect analysis of families affected by familial adenomatous polyposis (FAP) using polymorphic DNA markers

Family	*No of family members tested/total*	*No already known to be affected*	*No of family members at risk**	*Range of age of family members at risk*	*No of family members at risk found to be affected/not affected*
1	6/8	2	1	15	0/1
2	11/11	4	3	13–18	0/3
3	6/23	2	2	28–30	0/2
4	23/89	6	5	14–20	1/4
5	7/7	4	1	20	0/1
6	15/24	4	3	11–20	1/2
7	13/31	6	2	25–30	0/2
8	14/32	4	4	20–26	2/2
9	9/9	4	3	20–27	0/3
10	9/13	2	4	13–20	2/2
11	21/80	11	5	16–20	1/4
12	12/24	6	1	24	0/1
13	13/46	3	2	20–26	0/2
14	7/14	2	4	11–16	1/3
15	6/6	3	3	12–27	1/2
16	4/38	2	1	20	0/1
17	8/30	4	2	18–23	1/1
18	24/79	13	2	14–17	0/2
19	5/20	1	3	25–30	ND/3
20	4/10	1	2	13–14	ND/2
Total	217/594	84	53	—	10/43

*FAP not diagnosed, age under 30.

The polymorphic markers were APC 1458.5, 37AB, YN5.64, MBC, LNS, CB83 and CB26. No recombination events were detected. Figure 1 shows typical examples of the segregation of the disease genes in these families. In the first family (fig.1A) the haplotype associated with FAP was defined as CB26 (4), YN5.64 (5), CB83 (1), MBC (3), APC 1458.5 (2), based on the analysis of the affected members (II-6, II-8, III-1, III-4). Of the four presymptomatic, at risk individuals (III-2, III-3, III-5, III-6), two (III-3, III-6) appeared to have inherited the disease gene and to carry a risk of FAP of 99% (calculated as described in the Materials and methods section). Similarly, in family2 (fig.1B), the haplotype YN 5.64 (2), CB83 (1), MBC (3), APC 1458.5 (1), LNS (1), 37AB (2), shared by the affected members (II-3, III-1, III-4), was also found in one (III-5) of the three (III-2, III-3, III-5) presymptomatic relatives.

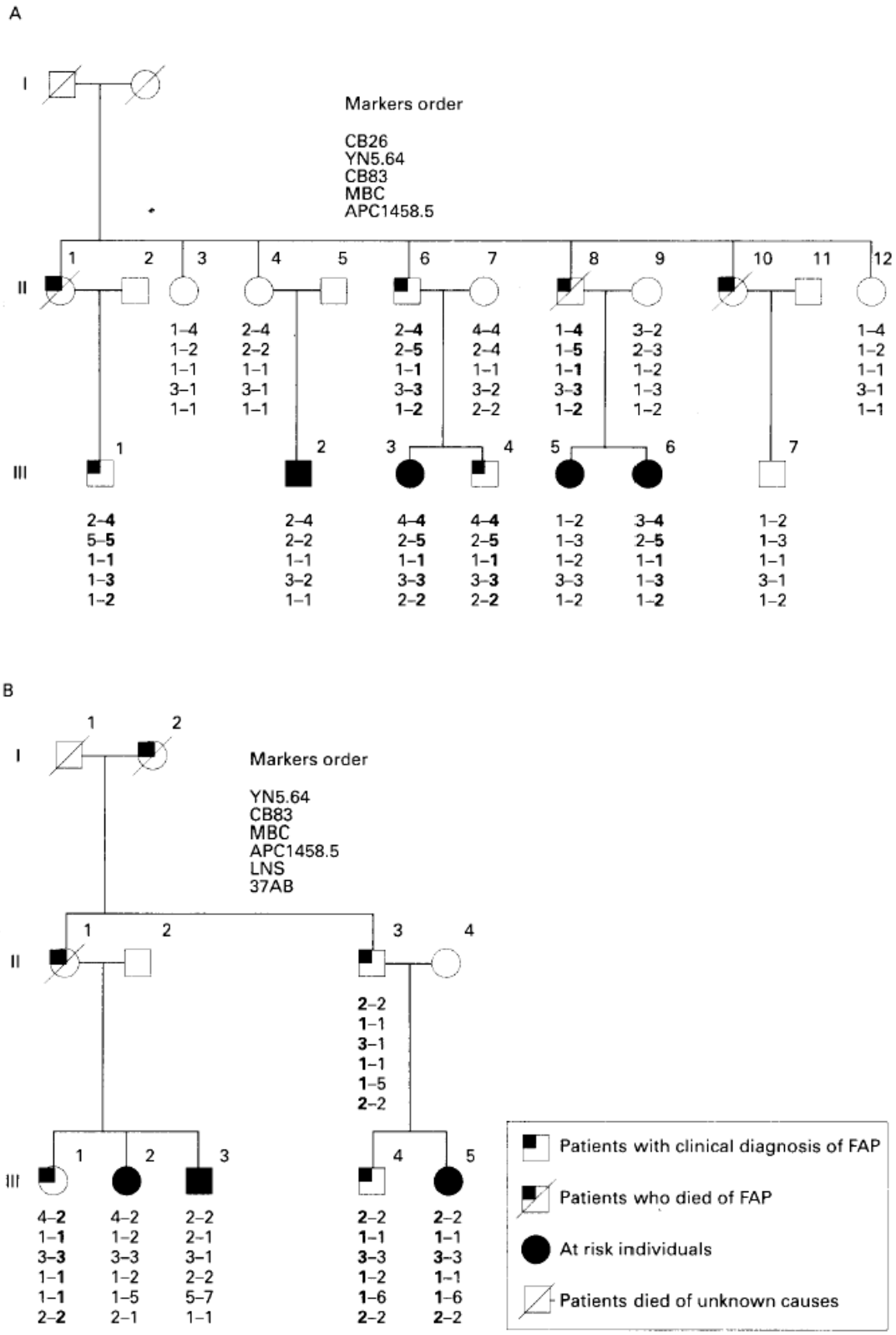
A
I
Markers order
CB26
YN5.64
CB83
MBC
APC1458.5
II
1 2 3 4 5 6 7 8 9 10 11 12
1–4 1–2 1–1 3–1 1–1
2–4 2–2 1–1 3–1 1–1
2–4 2–5 1–1 3–3 1–2
4–4 2–4 1–1 3–2 2–2
1–4 1–5 1–1 3–3 1–2
3–2 2–3 1–2 1–3 1–2
1–4 1–2 1–1 3–1 1–1
III
1 2 3 4 5 6 7
2–4 5–5 1–1 1–3 1–2
2–4 2–2 1–1 3–2 1–1
4–4 2–5 1–1 3–3 2–2
4–4 2–5 1–1 3–3 2–2
1–2 1–3 1–2 3–3 1–2
3–4 2–5 1–1 1–3 1–2
1–2 1–3 1–1 3–1 1–2
B
I
1 2
Markers order
YN5.64
CB83
MBC
APC1458.5
LNS
37AB
II
1 2 3 4
2–2 1–1 3–1 1–1 1–5 2–2
III
1 2 3 4 5
4–2 1–1 3–3 1–1 1–1 2–2
4–2 1–2 3–3 1–2 1–5 2–1
2–2 2–1 3–1 2–2 5–7 1–1
2–2 1–1 3–3 1–2 1–6 2–2
2–2 1–1 3–3 1–1 1–6 2–2
Patients with clinical diagnosis of FAP
Patients who died of FAP
At risk individuals
Patients died of unknown causes

Figure 1. Families with familial adenomatous polyposis pedigree: family N. 8 **A** and N. 15 **B** (see table). Segregation of the disease associated haplotype (shown in bold) is indicated. Chromosome are defined by the allele number of five or six markers which, starting from the top, include CB26, YN5.64, CB83, MBC, APC 1458.5 for family 1 and YN5.64, CB83, MBC, APC 1458.5, LNS, 37AB for family 2. The numbers
represent different sizes of DNA associated with each marker. The numbers are arranged as the markers are as if on chromosome 5 in and around the APC gene. The phase of the markers can be worked out by interpreting which marker came from each parent and grandparent. In this way the disease chromosome can be identified with the markers suggesting the affected status highlighted. The marker numbers themselves only predict affected status in a particular family.

Discussion

Since the first identification of the APC gene on chromosome 5, several polymorphic markers and mutations associated with FAP have been detected and used to analyze affected families and estimate the risk to presymptomatic individuals [20].

This approach has allowed for an early diagnosis of FAP, reducing the need for endoscopic screening and producing an overall improvement of the patients' quality of life [28]. We studied, using a panel of polymorphic markers, a collection of samples derived form 20 families who had been enrolled in the Italian Registry of Polyposis. These families were selected because previous direct tests had failed to identify a specific FAP mutation. In 90% of cases we identified the haplotype associated with the disease gene and followed its segregation, thus identifying the presymptomatic members who carried the mutation. YN5.64, LNS and APC1458.5 were the most informative markers, and no recombination events were ever detected even for the markers which were more distant from the APC locus [29].

To measure the risk to presymptomatic relatives we used the same approach, based on Bayes's theorem that is commonly applied to other genetic diseases, such as cystic fibrosis and Duchenne muscular dystrophy, and used in forensic medicine to calculate the probability of paternity [30, 31]. No conclusion could be reached in the analysis of two families because there was an insufficient number of probands or specimens. These cases, which sometimes limit the use of the indirect method, can be further evaluated by investigating the presence of specific mutations. A variety of gene amplification and detection techniques have in fact been developed, including SSCP, denaturing gradient gel electrophoresis, amplification refractory mutation systems, and the PTT [13-18]. However, the impossibility of screening the whole APC gene reduces the value of these mutation analyses, whose success rate is usually no better than 60-70% [19]. The indirect approach is still necessary in some families. Indeed, our data as well as those reported by other authors, [16 29 32-

[35] indicate that indirect analysis, when performed with a large number of informative markers, can be used as one of the first screening methods for families with FAP. Even when there are too few live probands, analysis of DNA obtained from paraffin embedded tissue from deceased patients may be possible. As quick results are needed and the sensitivity of the direct approach is variable, often requiring a lengthy analysis of several exons, it is suggested that linkage analysis should be used after a first unsuccessful direct test.

The combined application of direct mutation analysis techniques can further enhance the diagnostic efficacy of these procedures and contribute to a better management of patients with FAP.

Acknowledgements

The study was supported in part by a grant from A.I.R.C., contract CAP. 420.198.602 “Suscettibilita’ genetica al cancro colorettale” and from C.N.R. contract n. 95.00569.P39. I. Gazzoli was recipient of a fellowship from Lega Italiana per la Lotta Contro i Tumori.

References

1. Boland CR IS, Kim YS. Colonic plyps and the gastrointestinal polyposis syndrome. Gastrointestinal disease 1989; In Sleizinger MH, Fordtran JS, eds. Philadelphia Saunders.:1483.
2. Nakamura Y. The role of the adenomatous polyposis coli (APC) gene in human cancers. Adv Cancer Res 1993;62:65-87.
3. Reed TE NJ. A genetic study of multiple polyposis of the colon. Am. J. Hum. Genet 1995;7:236-63.
4. Murday V, Slack J. Inherited disorders associated with colorectal cancer. Cancer Surv 1989;8:139-57.
5. Leppert M, Dobbs M, Scambler P, O'Connell P, Nakamura Y, Stauffer D, Woodward S, Burt R, Hughes J, Gardner E, et al. The gene for familial polyposis coli maps to the long arm of chromosome 5. Science 1987;238:1411-3.
6. Bodmer WF, Bailey CJ, Bodmer J, Bussey HJ, Ellis A, Gorman P, Lucibello FC, Murday VA, Rider SH, Scambler P, et al. Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 1987;328:614-6.
7. Groden J, Thliveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, Joslyn G, Stevens J, Spirio L, Robertson M, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 1991;66:589-600.
8. Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A, Koyama K, Utsunomiya J, Baba S, Hedge P. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 1991;253:665-9.
9. Tops CM, Wijnen JT, Griffioen G, von Leeuwen IS, Vasen HF, den Hartog Jager FC, Breukel C, Nagengast FM, van der Klift HM, Lamers CB, et al. Presymptomatic diagnosis of familial adenomatous polyposis by bridging DNA markers. Lancet 1989;2:1361-3.
10. Dunlop MG, Wyllie AH, Steel CM, Piris J, Evans HJ. Linked DNA markers for presymptomatic diagnosis of familial adenomatous polyposis. Lancet 1991;337:313-6.
11. Petersen GM, Slack J, Nakamura Y. Screening guidelines and premorbid diagnosis of familial adenomatous polyposis using linkage. Gastroenterology 1991;100:1658-64.
12. Petersen GM, Francomano C, Kinzler K, Nakamura Y. Presymptomatic direct detection of adenomatous polyposis coli (APC) gene mutations in familial adenomatous polyposis. Hum Genet 1993;91:307-11.

13. Miyoshi Y, Ando H, Nagase H, Nishisho I, Horii A, Miki Y, Mori T, Utsunomiya J, Baba S, Petersen G, et al. Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients. Proc Natl Acad Sci U S A 1992;89:4452-6.
14. Nagase H, Miyoshi Y, Horii A, Aoki T, Petersen GM, Vogelstein B, Maher E, Ogawa M, Maruyama M, Utsunomiya J, et al. Screening for germ-line mutations in familial adenomatous polyposis patients: 61 new patients and a summary of 150 unrelated patients. Hum Mutat 1992;1:467-73.
15. Varesco L, Gismondi V, James R, Robertson M, Grammatico P, Groden J, Casarino L, De Benedetti L, Bafico A, Bertario L, et al. Identification of APC gene mutations in Italian adenomatous polyposis coli patients by PCR-SSCP analysis. Am J Hum Genet 1993;52:280-5.
16. Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, Hamilton SR, Vogelstein B, Kinzler KW. Molecular diagnosis of familial adenomatous polyposis. N Engl J Med 1993;329:1982-7.
17. van der Luijt R, Khan PM, Vasen H, van Leeuwen C, Tops C, Roest P, den Dunnen J, Fodde R. Rapid detection of translation-terminating mutations at the adenomatous polyposis coli (APC) gene by direct protein truncation test. Genomics 1994;20:1-4.
18. Prosser J, Condie A, Wright M, Horn JM, Fantes JA, Wyllie AH, Dunlop MG. APC mutation analysis by chemical cleavage of mismatch and a protein truncation assay in familial adenomatous polyposis. Br J Cancer 1994;70:841-6.
19. Meera Khan P MJ. APC mutations and linkage for pre-symptomatic diagnosis of familial adenomatous polyposis. Proceedings of the third EuroFAP worshop. Leiden; The Netherlands. Denmark: The Panum Institute, University of Copenhagen 1993:15-20.
20. Breukel C, Tops C, van Leeuwen C, van der Klift H, Fodde R, Khan PM. AT repeat polymorphism at the D5S122 locus tightly linked to adenomatous polyposis coli (APC). Nucleic Acids Res 1991;19:6665.
21. Breukel C, Tops C, van Leeuwen C, van der Klift H, Nakamura Y, Fodde R, Khan PM. CA repeat polymorphism at the D5S82 locus, proximal to adenomatous polyposis coli (APC). Nucleic Acids Res 1991;19:5804.
22. van Leeuwen C, Tops C, Breukel C, van der Klift H, Deaven L, Fodde R, Khan PM. CA repeat polymorphism within the MCC (mutated in colorectal cancer) gene. Nucleic Acids Res 1991;19:5805.
23. van Leeuwen C, Tops C, Breukel C, van der Klift H, Fodde R, Khan PM. CA repeat polymorphism at the D5S299 locus linked to adenomatous polyposis coli (APC). Nucleic Acids Res 1991;19:5805.
24. Spirio L, Joslyn G, Nelson L, Leppert M, White R. A CA repeat 30-70 KB downstream from the adenomatous polyposis coli (APC) gene. Nucleic Acids Res 1991;19:6348.
25. Joslyn G, Carlson M, Thliveris A, Albertsen H, Gelbert L, Samowitz W, Groden J, Stevens J, Spirio L, Robertson M, et al. Identification of deletion mutations and three new genes at the familial polyposis locus. Cell 1991;66:601-13.
26. Hoving E SSB, Baxter W eet al. Constant denaturant gel electrophoresis a modification of denaturing gradient gel electrophoresis: in mutation detection. Mutat Res 1991;262:63-71.
27. Schmitt. An elementary introduction to Bayesian statistics. 1964.
28. Faragher I VD. Genetic testing and screening in colon cancer. Diagnostic Oncology 1995;4:80-6.
29. Eckert WA, Jung C, Wolff G. Presymptomatic diagnosis in families with adenomatous polyposis using highly polymorphic dinucleotide CA repeat markers flanking the APC gene. J Med Genet 1994;31:442-7.
30. Friedman. The molecular pathology of Duchenne and Becker muscular dystrophy. Molecular pathology approaches to diagnosing human disease in the clinical laboratory. New York: Caroline Academic Press 1995:119-49.
31. Callanan. DNA analysis of inherited genetic changes: counselling and social issues.Molecular pathology approaches to diagnosing human disease in the clinical laboratory. New York: Caroline Academic Press 1995:171-91.
32. Park JG, Han HJ, Kang MS, Nakamura Y. Presymptomatic diagnosis of familial adenomatous polyposis coli. Dis Colon Rectum 1994;37:700-7.
33. Bapat BV, Parker JA, Berk T, Cohen Z, McLeod RS, Ray PN, Stern HS. Combined use of molecular and biomarkers for presymptomatic carrier risk assessment in familial adenomatous polyposis: implications for screening guidelines. Dis Colon Rectum 1994;37:165-71.
34. Caspari R, Friedl W, Boker T, Augustin A, Mandl M, Jaeger K, Gallkowski K, Propping P. Predictive diagnosis in familial adenomatous polyposis: evaluation of molecular genetic and ophthalmologic methods. Z Gastroenterol 1993;31:646-52.
35. MacDonald F, Morton DG, Rindl PM, Haydon J, Cullen R, Gibson J, Neoptolemos JP, Keighley MR, McKeown CM, Hulten M. Predictive diagnosis of familial adenomatous polyposis with linked DNA markers: population based study. Bmj 1992;304:869-72.

Chapter 3

Microsatellite instability in colorectal-cancer patients with suspected genetic predisposition.

Calistri D, Presciuttini S, Buonsanti G, Radice P, Gazzoli I, Pensotti V, Sala P, Eboli M, Andreola S, Russo A, Pierotti M, Bertario L and Ranzani GN.

International Journal of Cancer 2000 Jan 20; 89(1):87-91.

Abstract

Hereditary non-polyposis colorectal cancer (HNPCC) is a dominantly inherited syndrome linked to DNA-*mismatch-repair* (MMR) gene defects, which also account for microsatellite instability (MSI) in tumor tissues. Diagnosis is based mainly on family history, according to widely accepted criteria (Amsterdam Criteria: AC). Aim of this work was to assess MSI in colorectal-cancer patients with suspected genetic predisposition, and to verify whether MSI represents a tool to manage MMR gene (*hMSH*2 and *hMLH*1) mutation analysis. We investigated 13 microsatellites (including the 5 NCI/ICGHNPCC markers) in 45 patients with suspected hereditary predisposition (including 16 subjects from HNPCC families fulfilling the AC). We found MSI-H (high frequency of instability, *i.e.,* in >30% of the markers) in 85% of the HNPCC patients and in 16% of the non-HNPCC subjects. The 5 NCI/ ICG-HNPCC microsatellites proved to be the most effective in detecting MSI, being mononucleotide repeats the most unstable markers. We investigated the association between *hMSH*2- and *hMLH*1 gene mutations and MSI. Our results indicate that AC are highly predictive both of tumor instability and of MMR-gene mutations. Therefore, as the most likely mutation carriers, HNPCC subjects might be directly analyzed for gene mutations, while to test for MSI in selected non-HNPCC patients and to further investigate MMR genes in MSI-H cases, appears to be a cost-effective way to identify subjects, other than those from kindred fulfilling AC, who might benefit from genetic testing.

Introduction

Hereditary non-polyposis colorectal cancer (HNPCC) is an autosomal dominant disease characterized by early onset of single or multiple (synchronous or methacronous) colorectal cancers (CRC). Extracolonic tumors, which predominantly affect endometrium, stomach, small intestine and the hepatobiliary and urinary tracts, can also cluster in HNPCC families. To date, the syndrome has been defined on the exclusive basis of family history. A widely accepted definition based on the "Amsterdam Criteria" (AC) requires that at least 3 individuals contract histologically verified colorectal cancers in 2 consecutive generations, that one subject be a first-degree relative of the others, and that at least one of them develops a cancer before reaching the age of 50 (Vasen et al., 1991).

The syndrome is associated with genes involved in DNA-nucleotide- mismatch repair (MMR). Germline mutations in 5 MMR genes (hMSH2, hMLH1, hPMS1, hPMS2 and hMSH6) have been identified in affected members of HNPCC families; the great majority of mutations are found in either hMSH2 or hMLH1 genes (quoted by Peltomäki et al., 1997).

MMR genes play a crucial role in DNA-replication fidelity: their dysfunction has been found to be associated both with increased mutation rate (Bhattacharyya et al., 1994) and with widespread genome instability (Aaltonen et al., 1993). Genetic instability manifests as alterations in the size of simple repeated sequences (microsatellites) in tumor DNA as compared with control DNA of the same patient (Aaltonen et al., 1993). The resulting tumor phenotype, commonly called MSI, from "microsatellite instability," or RER1, from "replication error," is almost always present (occurrence rate .70%) in the cancer tissues of HNPCC patients (Aaltonen et al., 1993). MSI also characterizes a fraction of sporadic colon cancers (Ionov et al., 1993; Thibodeau et al., 1993) as well as a fraction of other sporadic tumors (Peltomäki et al., 1993).
Germline mutations in the above-mentioned genes have been reported to account for approximately 25 to 70% of kindred with HNPCC fulfilling the AC (Weber et al., 1997; Liu et al., 1996; quoted by Peltomäki et al., 1997). On the other hand, germline mutations have been found in families that do not meet the AC (Nyström-Lahti et al., 1996; Moslein et al., 1996; Wijnen et al., 1997; Genuardi et al., 1998) and in patients with early onset of colorectal cancer and no family history of cancer (Liu et al., 1995).

The aim of this study was to investigate tumor instability and to verify whether MSI represents a tool with which to manage MMR-gene mutation analysis in patients with suspected hereditary predisposition to colorectal cancer. For this purpose, we analyzed MSI in a panel of patients that consisted of affected members of HNPCC families fulfilling the AC and patients who met at least one of the following criteria: family history of cancer, early onset of cancer, development of multiple tumors. Some subjects of this panel had been investigated previously for hMSH2 and hMLH1 gene mutations (Pensotti et al., 1997); gene screening was extended to other subjects to investigate the association between gene mutations and MSI. This study indicates that microsatellite assessment may be of use to direct gene-mutation analysis, the results of which are crucial for counseling and management of colorectal-cancer families.

Materials and methods

Patients. We analyzed 45 unrelated CRC patients classified as follows. HNPCC, individuals from families fulfilling the stringent AC for HNPCC diagnosis (13 cases); HNPCC*, individuals from HNPCC families partially documented (3 cases); HNPCC-modified, individuals from families meeting 2 out of 3 AC for HNPCC diagnosis (4 cases); family-history, subjects having at least one first-degree relative affected with colorectal cancer (17 cases); juvenile, subjects less than 50 years old at diagnosis and no family history of cancer (7 cases); multiple neoplasms, subjects with multiple tumors and no family history of cancer (1 case). Among non- HNPCC patients, 2 HNPCC*, 1 HNPCC-modified and 3 family history subjects were characterized by the development of multiple tumors.

Tissue samples, microsatellite and gene analysis. Tissue samples from colon cancers were taken from the files of the Division of Pathology, Istituto Nazionale Tumori, Milan. Both tumor and non-tumor tissues of each patient were obtained from paraffin blocks of formalin-fixed surgical specimens or from frozen tissue fragments. Neoplastic samples contained at least 35 to 40% tumor cells, as evaluated by microscopic analysis. DNA was isolated as described (Ranzani et al., 1995).

We analyzed instability by screening 13 microsatellites, including 3 mono-, 9 di- and one tetranucleotide repeats. These were: BAT-25, BAT-26, BAT-40; D2S123, D5S346, D17S250, D2S177, D3S1076, D5S433, D11S904, D17S796, D18S59; and HUMTH01. BAT-25, BAT-26, BAT-40, D2S123, D5S346 and D17S250 loci were amplified with the specific primers reported by

Dietmaier et al., (1997) and analyzed following the conditions reported by the same authors. For the remaining loci, PCR reactions were carried out with the specific primers reported by Jones et al. (1992) and in The Généthon Microsatellite Map Catalogue (Généthon, Human Genome Research Centre, Evry, France) in a 25-µl volume containing 50 ng of template DNA, 0.5 µM each unlabelled primer, 200 µM dGTP, dTTP, dCTP, 25 µM dATP, 2.5 µCi (α-^{35}S) dATP, 1 mM $MgCl_2$, 0.2 mg/ml BSA, 50 mM KCl, 10 mM Tris-HCl, pH 8.4, and 1 unit of Taq polymerase. Samples were processed through 27 to 30 cycles, each cycle consisting of 1 min at 94°C, 2 min at 58°C and 1 min and 10 sec at 72°C.

PCR products from tumor and corresponding control DNA were loaded in parallel on standard 6% polyacrylamide, 7 M urea DNA-denaturing sequencing gels, and visualized by autoradiography. The autoradiogram analysis was carried out by 2 independent observers. All experiments in which extra bands were detected in tumor samples were entirely repeated a further 2 times; the results were consistent in all the tests performed.

DNA for hMSH2 and hMLH1 gene analysis was obtained from peripheral-blood samples of the same subjects in which the MSI analysis was performed. Mutation screening was carried out via exon-by-exon SSCP analysis on 13 samples as reported (Pensotti et al., 1997).

Statistical analysis. An unconditional multiple logistic-regression model was used to compare the association between selected characteristics and MSI status. Odds ratios (OR) and corresponding 95% confidence intervals (CI) were computed (Breslow and Day, 1980). All OR were adjusted for age and sex.

Results

We investigated MSI in 45 CRC patients with suspected hereditary predisposition, including 13 subjects from families fulfilling the stringent AC for HNPCC diagnosis, and 32 non-HNPCC subjects. To assess MSI, we analyzed 13 markers. Out of 45 tumors, 20 showed instability at one or more loci. Some studies indicate that a 30–40% cutoff value of unstable markers, out of the total number of markers tested, can define 2 discrete groups of tumors, MSI-L (low frequency of MSI) and MSI-H (high frequency of MSI), only one of which (MSI-H) appears to result from a defective MMR system (Olschwang *et al.,* 1997; Dietmaier *et al.,* 1997). By using a cutoff value of ≥30%,

we classified 16 samples as MSI-H and 4 samples as MSI-L. The remaining tumors showed 0/13 unstable markers and were classified as MSS (microsatellite- stable tumors). Figure 1 shows instability at different loci. The microsatellite panel we utilized includes the 5 MSI diagnostic loci designated by the National Cancer Institute (NCI)/ International Collaborative Group for HNPCC (ICG-HNPCC) (Boland *et al.,* 1998), according to whose recommendations tumors can be classified as MSI-H if 2 or more of these markers show instability.

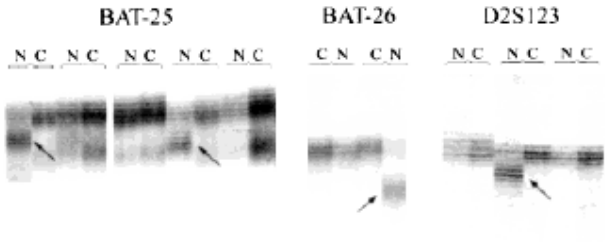

Figure 1. Genetic instability at different loci in colorectal cancer patients. PCR products from neoplasia (N) and control (C) DNA were loaded in parallel; arrows indicate additional bands in tumor samples as compared with corresponding control samples.

On the basis of these criteria, we identified 15 MSI-H and 3 MSI-L tumors. Results obtained with all investigated markers and with the NCI/ICG-HNPCC panel are compared in Table I.

TABLE I – COLORECTAL CANCERS SHOWING INSTABILITY

Tumour samples	NCI/ICG-HNPCC marker panel					NCI/ICG-HNPCC marker panel: unstable loci/ classifiable loci	All investigated markers: unstable loci/ classifiable loci
	BAT 25	BAT 26	D2S123	D5S346	D17S250		
1	+	+	−	−	−	**2/5**	**4/13**
2	+	+	+	−	+	**4/5**	**8/13**
3	+	+	+	−	+	**4/5**	**8/13**
4	−	+	−	−	−	1/5	**4/13**
5	+	+	−	−	+	**3/5**	**5/13**
6	+	+	+	n.c.	n.c.	**3/3**	**6/8**
7	+	+	+	+	+	**5/5**	**9/13**
8	+	+	n.c.	+	n.c.	**3/3**	**6/11**
9	+	−	−	−	+	**2/5**	**4/13**
10	+	+	+	−	+	**4/5**	**7/12**
12	+	+	−	−	+	**3/5**	**5/13**
13	+	+	+	−	−	**3/5**	**7/13**
14	+	+	+	−	−	**3/5**	**5/13**
18	−	−	−	−	−	0/5	1/13
19	+	+	+	−	+	**4/5**	**8/12**
21	−	−	−	−	−	0/5	1/10
27	−	−	+	−	−	1/5	1/13
31	−	+	−	−	−	1/5	2/13
38	+	+	−	+	+	**4/5**	**8/13**
43	+	+	+	−	+	**4/5**	**10/13**

MSI-H tumour samples shown in bold type. n.c., not classifiable.

Among the MSI-H tumors, 11 were from HNPCC patients (11/13, 85%) and 5 from non-HNPCC patients (5/32, 16%); all MSI-L tumors (4 cases) were from non-HNPCC subjects.

*hMSH*2- and *hMLH*1-germline-mutation analysis was available for 24 patients, including all HNPCC patients and non-HNPCC subjects with a strong family history of cancer and/or high tumor instability; no material was available for case 27 (Table II). A mutation in the *hMSH*2 or the *hMLH*1 gene was found in 8 individuals; 6 mutations have been described earlier (Pensotti *et al.,* 1997) and 2 were identified during the present screening (Table I). Of the mutation carriers, 6 were from families fulfilling the AC (6/13, 46%), one was from an HNPCC family partially documented, and one was a subject of young age with family history of cancer. All mutation carriers were characterized by high tumor instability. No detectable mutations were present in the remaining subjects with MSI-H tumors, *i.e.,* 5 HNPCC and 3 non-HNPCC patients. Table II summarizes all data, and shows, for each patient: features suggestive of cancer predisposition, MSI, *hMSH*2 and *hMLH*1 gene status, tumor site and stage, and age at diagnosis.

TABLE II – PATIENT CLASSIFICATION, TUMOUR CHARACTERISTICS, MSI AND MMR GENE MUTATIONS

Sample	Patient classification[1]	MSI[2]	Mutated gene	Nucleotide change	Protein change/codon affected	Tumour site[3]	Dukes' stage	Age at diagnosis
1	HNPCC	H	n.d.			P	B	41
2	HNPCC	H	*hMLH1*[4]	1852 del AAG	del K/618	P	B	45
3	HNPCC	H	*hMLH1*[5]	IVS6 + 3 A > G	Exon skipping out of frame	P	B	45
4	HNPCC	H	*hMSH2*	278 del TT	Frameshift/Stop at 98	D	A	31
5	HNPCC	H	*hMSH2*[5]	IVS5 + 3 A > T	Exon skipping in frame	D	B	47
6	HNPCC	H	n.d.[5]			P	B	66
7	HNPCC	H	n.d.[5]			D	C	49
8	HNPCC	H	n.d.[5]			P	B	62
9	HNPCC	H	*hMLH1*[5]	382 G > C	A > P/128	P	B	32
12	HNPCC	H	n.d.			P	C	39
13	HNPCC	H	*hMSH2*[5]	IVS5 + 3 A > T	Exon skipping in frame	D	C	30
16	HNPCC	S	n.d.[5]			D	A	47
20	HNPCC	S	n.d.			D	B	57
11	HNPCC*	S	n.d.[5]			D	B	43
14	HNPCC*	H	*hMLH1*[5]	731 G > A	G > D/244	P	B	60
23	HNPCC*	S	n.d.			D	B	58
10	HNPCC mod	H	n.d.			D	B	37
18	HNPCC mod	L	n.d.			P	B	43
21	HNPCC mod	L	n.d.			D	B	73
27	HNPCC mod	L	—			D	C	57
15	fam his	S	n.d.			D	B	37
17	fam his	S	n.d.			P	D	61
19	fam his	H	n.d.			P	B	81
22	fam his	S	—			D	B	81
26	fam his	S	—			D	B	55
28	fam his	S	—			D	D	59
29	fam his	S	—			D	B	86
30	fam his	S	—			P	C	71
31	fam his	L	—			P	A	65
32	fam his	S	—			D	B	59
34	fam his	S	—			P	D	74
36	fam his	S	—			P	D	56
37	fam his	S	—			D	B	78
39	fam his	S	—			D	A	70
42	fam his	S	—			D	C	59
43	fam his	H	*hMSH2*	1435 A > T	S > C/479	D	A	30
44	fam his	S	—			D	C	79
24	juv	S	—			P	C	38
25	juv	S	—			D	D	42
33	juv	S	—			P	C	41
35	juv	S	—			D	C	43
40	juv	S	—			D	A	39
41	juv	S	—			D	D	32
45	juv	S	—			P	B	49
38	multiple tum	H	n.d.			P	B	64

[1]HNPCC, individuals from families fulfilling the stringent AC; HNPCC*, individuals from HNPCC families only partially documented; HNPCC mod, individuals from families meeting 2 out of 3 AC; fam his, subjects having at least one first-degree relative affected with colorectal cancer; juv, subjects less than 50 years old at diagnosis and no family history of cancer; multiple tum, subjects with multiple tumors and no family history of cancer. -[3]H, high instability; L, low instability; S, no instability. -[3]P, proximal to splenic flexure; D, distal to splenic flexure. -[4]Wijnen et al. (1997). -[5]Pensotti et al. (1997). – n.d., no hMSH2 or hMLH1 gene mutations detected.

Table III shows tumor instability in relation to patient characteristics. MSI-H status was significantly associated with increased risk of tumor location in proximal colon (OR 7.5, 95% CI 1.5–37.4) and with HNPCC (OR 16.6, 95% CI 1.4 –203.1). The proportion of MSI-H cases did not differ significantly according to other variables, although a weak positive association was evident for tumors of Dukes' early stages (81% in MSI-H *vs.* 55% in MSS plus MSI-L).

TABLE III – PATIENT AND TUMOUR CHARACTERISTICS IN RELATION TO TUMOUR INSTABILITY

Patient and tumour characteristics	MSS + MSI-L tumours	MSI-H tumours	Total	OR (95% CI)
Sex				
Males	15 (51.7)	10 (62.5)	25 (55.6)	1
Females	14 (48.3)	6 (37.5)	20 (44.4)	0.7 (0.5–1.1)
Age (years)				
≥50	18 (62.1)	5 (31.3)	23 (51.1)	1
<50	11 (37.9)	11 (68.8)	22 (48.9)	4.0 (1.1–15.1)
Site				
Distal	20 (69.0)	6 (37.5)	26 (57.8)	1
Proximal	9 (31.0)	10 (62.5)	19 (42.2)	7.5 (1.5–37.4)
Dukes' stage				
C/D	13 (44.8)	3 (18.8)	16 (35.6)	1
A/B	16 (55.2)	13 (81.3)	29 (64.4)	4.6 (0.9–22.3)
Cancer predisposition				
Juvenile	7 (24.1)	1 (6.3)	8 (17.8)	1
Family history	15 (51.7)	2 (12.5)	17 (37.8)	1.4 (0.1–36.1)
HNPCC[1]	7 (24.1)	13 (81.3)	20 (44.4)	16.6 (1.4–203.1)
Mutation				
Unknown	21	—	21	
Negative	8	8	16	
Positive	—	8	8	

[1]Patients from HNPCC, HNPCC*, and HNPCC-mod families (see "Material and Methods") have been pooled.

Discussion

Each of the 13 markers we analyzed manifested instability in different tumor samples, allowing detection of 16 MSI-H cases. However, the 5 NCI/ICG-HNPCC markers proved to be the most effective in detecting instability. Indeed, by using only these microsatellites, and following the NCI/ICG-HNPCC recommendations (Boland *et al.,* 1998), we would have classified as MSI-H 15 out of 16 cases (Table I).

MSI-H instability proved to characterize 85% of the HNPCC patients meeting the AC, and 16% of the remaining cases. The frequency we found for the HNPCC patients was similar to those reported earlier (63–92%; Liu *et al.,* 1996; Jass *et al.,* 1995). The frequency we found for the remaining cases was slightly higher than those reported for most patients with sporadic colorectal cancer (12–28%; Ionov *et al.,* 1993; Thibodeau *et al.,* 1993).

In agreement with other findings (Aaltonen *et al.,* 1993; Lothe *et al.,* 1993), MSI-H status proved to be significantly associated with increased risk of tumour location in proximal colon and with HNPCC (Table III).

We identified 8 mutation carriers, in *hMSH*2 or in *hMLH*1 genes. All were characterized by MSI-H tumors. Gene alterations included intragenic deletions, frameshift, non-sense and mis-sense mutations. The latter were considered as "bona fide" pathogenetic, since segregating with the disease in the family and not found in control individuals (Pensotti *et al.,* 1997), or showing a dominant mutator effect in yeast assay (Shimodaira *et al.,* 1998). Mutations were present in 46% (6/13) of the patients from HNPCC families (Table II). Similar frequencies have been reported for patients from kindred fully meeting the AC (quoted by Peltomäki *et al.,* 1997). No germline mutations in *hMSH*2 or *hMLH*1 genes or tumor instability were detected in 2 HNPCC patients (16 and 20, Table II). This observation suggests either that some cancers are due to germline mutations in genes other than those involved in DNA MMR, or that a minority of families selected on the basis of the AC represents a non-genetic clustering of patients.

Data reported in Table II clearly demonstrate that AC are highly predictive both of tumor instability and of germline mutations in DNA MMR genes. Due to the high frequency of instability in HNPCC patients (85%), microsatellite investigation appears relatively unimportant for directing genetic analysis in patients from kindred strictly complying with the AC. Therefore, as the most likely mutation carriers, these individuals might be directly analyzed for hMSH2 and hMLH1 gene mutations.

Among the non-HNPCC patients, only 5 of 32 individuals showed tumor instability. We analyzed MMR genes in 11 subjects characterized by strong family history of cancer and/or tumor instability and identified a disease-causing mutation in 2 of the 5 MSI-H subjects. Due to the low frequency of instability (16%) in non-HNPCC patients, characterization of tumor phenotype prior to genetic testing might be of great relevance in these patients. Indeed, direct genetic analysis by any of the

methods commonly utilized for germline-mutation screening is much more expensive and time- and labor-consuming than MSI assessment by diagnostic microsatellites and further investigation of MMR genes in subjects showing instability.

In the whole sample of non-HNPCC patients, the mutation frequency was 6% (2/32), a value in good agreement with reports that patients from families not fulfilling the AC show very low frequencies (about 10%) of MMR-gene mutations (Nyström-Lahti et al., 1996; Moslein et al., 1996; Wijnen et al., 1997; Genuardi et al., 1998).

Aaltonen et al. (1998) have reported that investigation of MSI on unselected CRC patients can be used to identify MMR-gene mutation carriers. However, to improve the efficiency of screening for mutation carriers, they recommended testing for MSI in selected patients showing at least one of the hallmarks of HNPCC (family history, young age of onset, development of multiple tumors) and to further investigate MMR genes in patients found to have tumor instability. Our approach is in line with this scheme which, in our opinion, represents a cost-effective way of identifying patients, other than those from kindred fulfilling AC that might benefit from MMR-gene testing. Wijnen et al. (1998) used clinical findings to devise a logistic model for estimating the likelihood of a MMR-gene mutation in families with clustering of CRC.

It is worth stressing that all patients we investigated for MSI had been selected on the basis of family history of cancer or other features suggestive of cancer predisposition. Consequently, we do not consider the absence of MSI as an absolute criterion for eliminating patients from other possible genetic analyses. It is possible that mutations in MMR genes other than hMSH2 and hMLH1 might be associated with a mild phenotypic effect in cancer cells. Moreover, patients without instability could carry germline mutations in genes other than those involved in DNA MMR. The finding of TGFß-RII-germline mutations in patients without tumor instability and with family history of cancer not in accordance with the AC (Lu et al., 1998) suggests that alterations of genetic pathways different from the mutator pathway may give rise to a predisposition to non-polyposis colorectal cancer.

Acknowledgements

G.B. was supported by an AIRC fellowship. We thank Dr. E. Arbustini, IRCCS Policlinico S. Matteo-Pavia, for kindly providing DNA-sequencing facilities.

References

Aaltonen L.A. and 13 Others. Clues to the pathogenesis of familial colorectal cancer. Science, 260**,** 812–816 (1993).

Aaltonen L.A., Salovaara R., Kristo P., Canzian,F., Hemminki A., PeltomakiI,P., Chadwick R.B., Kaariainen H., Eskelinen M., Jarvinen H., Mecklin J.P. and de la Chapelle. A. Incidence of hereditary non-polyposis colorectal cancer and the feasibility of molecular screening for the disease. New Engl. J. Med., 338, 1481–1487 (1998).

Bhattacharyya N.P., Skandalis A., Ganesh A., Groden J. and Meuth M. Mutator phenotypes in human colorectal-carcinoma cell lines. Proc. nat. Acad. Sci. (Wash.), 91, 6319–6326 (1994).

Boland C.R., Thibodeau S.N., Hamilton S.R., Sidransky D., Eshleman J.R., Burt, R.W., Meltzer S.J., Rodriguez-Bigas M.A., Fodde R., Ranzani G.N. and 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., 58, 5248–5257 (1998).

Breslow N.E. and Day, N.E. Statistical methods in cancer research. The analysis of case-control studies, Vol. I, IARC Scientific Publication 32, IARC, Lyon (1980).

Dietmaier W., Wallinger S., Bocker T., Kullmann F., Fishel R. And Ruschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch-repair-protein expression. Cancer Res., 57**,** 4749–4756 (1997).

Genuardi M. and 18 Others. *MLH1* and *MSH2* constitutional mutations in colorectal cancer families not meeting the standard criteria for hereditary non-polyposis colorectal cancer. Int. J. Cancer, 75, 835–839 (1998).

Ionov Y., Peinado M.A., Malkhosya S., Shibata D. and Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature (Lond.), 363, 558–561 (1993).

Jass J.R., Cottier D.S., Jeevaratnam P., Pokos V., Holdaway K.M., Bowden M.L., van de Water N.S. and Browett P.J. Diagnostic use of microsatellite instability in hereditary non-polyposis colorectal cancer. Lancet, 346, 1200–1201 (1995).

Jones M.H., Yamakawa K. and Nakamura Y. Isolation and characterization of 19 dinucleotide-repeat polymorphisms on chromosome 3p. Hum. Mol. Genet., 1, 131–133 (1992).

Liu B., Farrington S.M., Petersen G.M., Hamilton,S.R., Parsons R., Papadopoulos N., Fujiwara T., Jen J., Kinzeler K.W., Wyllie A.H., Vogelstein B. and Dunlop M.G. Genetic instability occurs in the majority of young patients with colorectal cancer. Nature Med., 1, 348– 352 (1995).

Liu B. and 13 Others. Analysis of mismatch-repair genes in hereditary non- polyposis-colorectal-cancer patients. Nature Med., 2, 169–174 (1996).

Lothe R.A. and 15 Others. Genomic instability in colorectal cancer: relationship to clinic pathological variables and family history. Cancer Res., 53, 5849–5852 (1993).

Lu S-L., Kawabata M., Imamura T., Akiyama Y., Nomizu T., Miyazono K. and Yuasa Y. HNPCC associated with germline mutation in the TGF-b-type-II-receptor gene. Nature Genet., 19, 17–18 (1998).

Moslein G., Tester D.J., Lindor N.M., Honchel R., Cuningham J.M., French A.J., Halling K.C., Schwab M., Goretzki P. and Thibodeau S.N. Microsatellite instability and mutation analysis of *hMSH*2 and *hMLH*1 in patients with sporadic, familial and hereditary colorectal cancer. Hum. Mol. Genet., 5, 1245–1252 (1996).

Nystrom-Lahti M., Wu Y., Moisio A-L., Hofstra R.M.W., Odinga J., Mecklin J-P., Jarvinen H.J., Leisti J., Buys C.H.C.M., de la Chapelle A. and Peltomaki P. DNA mismatch-repair-gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum. Mol. Genet., 5, 763–769 (1996).

Oolschwang S., Hamelin R., Laurent-Puig P., Thuille B., De Rycke Y., Li Y-J., Muzeau F., Girodet J., Salmon R-J. and Thomas G. Alternative genetic pathways in colorectal carcinogenesis. Proc. Nat. Acad. Sci. (Wash.), 94, 12122–12127 (1997).

Peltomaki P. and 12 Others. Microsatellite instability is associated with tumours that characterize the hereditary non-polyposis-colorectal-carcinoma syndrome. Cancer Res., 53, 5853–5855 (1993).

Peltomaki P., Vase, H.F.A. and The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Mutations predisposing to hereditary non-polyposis colorectal cancer: database and results of a collaborative study. Gastroenterology, 113, 1146–1158 (1997).

Pensotti V. and 12 Others. Mean age of tumour onset in hereditary-nonpolyposis- colorectal-cancer (HNPCC) families correlates with the presence of mutations in DNA mismatch-repair genes. Genes Chromosomes Cancer, 19, 135–142 (1997).

Ranzani G.N., Luinetti O., Padovani L.S., Calistri D., Renault, B., Burrel M., AmadoriI, D., Fiocca R. and Solcia E. *p53*-gene mutations and protein nuclear accumulation are early events in intestinal-type gastric cancer but late events in diffuse type. Cancer Epidemiol. Biomarkers Prev., 4, 223–231 (1995).

Shimodaira H., Filosi N., Shibata H., Suzuki T., Radice P., KanamaruANAMARU, R., FriendS.H., Kolodner R.D. and Ihioka C. Functional analysis of human *MLH*1 mutations in *Saccharomyces cerevisiae.* Nature Genet., 19, 384–389 (1998).

Thibodeau S.N., BrenG. and SchaidD. Microsatellite instability in cancer of the proximal colon. Science, 260, 816–819 (1993).

Vasen H.F.A., Mecklin J-P., Meera Khan P. and LYNCH, H.T. The International Collaborative Group on Hereditary Non-polyposis Colorectal Cancer. Dis. Colon Rect., 34, 424–425 (1991).

Weber T.K., and 12 Others. Genomic DNA-based *hMSH*2 and *hMLH*1 mutation screening in 32 Eastern United States hereditary-non-polyposiscolorectal- cancer pedigrees. Cancer Res., 57, 3798–3803 (1997).

Wijnen J., Khan P.M., Vasen H., van der Klift, H., Mulder A, Moller, A., van Leeuwen-Cornelisse, I., BakkerB., Losekoot M., Molle P. and Fodde R. Hereditary-non-polyposis-colorectal-cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations. Amer. J. Hum. Genet., 61, 329–335 (1997).

Wijnen, J.T., Vasen H.F.A., Meera Khan P., Zwinderman A.H., van der Klift, H., Mulder A., Tops C., Molle P. and Fodde R. Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. New Engl. J. Med., 339, 511–518 (1998).

Chapter 4

Mean age of tumor onset in hereditary nonpolyposis colorectal cancer (HNPCC) families correlates with the presence of mutations in DNA mismatch repair genes.

Pensotti V, Radice P, Presciuttini S, Calistri D, Gazzoli I, Grimalt Perez A, Mondini P, Buonsanti G, Sala P, Rossetti C, Ranzani GN, Bertario L and Pierotti MA.

Genes, Chromosomes & Cancer 1997 Jul; 19(3):135-42.

Abstract

Fourteen Italian families affected with hereditary nonpolyposis colorectal cancer (HNPCC) were screened for germline mutations at three DNA mismatch repair (MMR) genes, *MSH2*, *MLH1*, and *GTBP*, by using a combination of different methods that included an in vitro synthesized protein assay, single-strand conformation polymorphism analysis, and direct sequencing. DNA alterations were observed in six instances, including a single base deletion in *MSH2* exon 14, an A-to-G transition in the splice donor site of *MLH1* exon 6, and two missense mutations in *MLH1* exons 5 and 9. A previously reported common mutation affecting the splice donor site of *MSH2* exon 5 was identified in two families. No mutations were detected in the *GTBP* gene. In total, eight of 16 Italian HNPCC families (50%), including two previously reported kindreds, were found to carry a mutation in MMR genes. We compared the mean age of colorectal cancer onset in the index cases (three patients for each family) between the two groups of kindreds, those with identified mutation vs. those without, and found that the first had a significantly lower value (43.0 vs. 53.7 years, P = 0.014). This finding suggests that HNPCC families with a more advanced age of tumor onset are less likely to be associated with known MMR genes.

Introduction

Hereditary nonpolyposis colorectal cancer (HNPCC) is an autosomal dominant genetic disease characterized by early occurrence of colorectal cancers and increased risk for neoplasms of other organs, including tumors of the endometrium, stomach, small intestine, hepatobiliary system, kidney, ureter, and ovary (Lynch et al., 1993). To provide a basis of uniformity in the ascertainment of HNPCC families, clinical guidelines have been proposed. These ‘‘Amsterdam criteria’’ are as follows: (1) histologically verified colorectal cancer in at least three relatives, one of whom must be a first-degree relative of the other two; (2) colorectal cancer involving at least two generations; and (3) at least one colorectal cancer case diagnosed before 50 years of age (Vasen et al., 1991). According to recent estimates, the proportion of HNPCC cases fulfilling these criteria corresponds to approximately 1% of all colorectal cancers (Kee and Collins, 1992; Aaltonen et al., 1994).

Four different genes have been found to be mutated in HNPCC patients, including *MSH2* in chromosome band 2p16 (Fishel et al., 1993; Leach et al., 1993), *MLH1* in 3p21 (Bronner et al., 1994; Papadopoulos et al., 1994), and *PMS1* and *PMS2* in 2q31–33 and 7p22, respectively (Nicolaides et al., 1994). These genes are the human counterparts of *Escherichia coli* and yeast genes whose products selectively bind DNA mismatches, initiating the repair process (Lu et al., 1993; Strand et al., 1993). Therefore, they are usually referred to as DNA mismatch repair (MMR) genes. A fifth MMR gene, termed *GTBP* (G/T binding protein), was recently identified (Drummond et al., 1995; Palombo et al., 1995). *GTBP* maps to chromosome band 2p16 within 0.5 megabase from *MSH2* and codes for a 160-kDa protein that interacts with the product of *MSH2*. Although mutations of the *GTBP* gene have been observed in colon carcinoma cell lines, germline alterations have so far not been reported in HNPCC patients (Papadopoulos et al., 1995; Liu et al.,1996).

The majority of germline mutations at the *MSH2* locus lead to the truncation of the protein product and appear to be dispersed along the entire coding region (Liu et al., 1994, 1996; Buerstedde et al., 1995; Wijnen et al., 1995; Lu et al., 1996; Maliaka et al., 1996; Moslein et al., 1996). The mutation spectrum of the *MLH1* gene is more heterogeneous and includes missense mutations, frame shifts, and deletions of entire exons due to substitutions at the splice donor or acceptor sites (Buerstedde et al., 1995; Han et al., 1995; Tannergard et al., 1995; Liu et al., 1996; Maliaka et al., 1996; Moslein

et al., 1996; Nystrom-Lahti et al., 1996; Wijnen et al., 1996). So far, only few germline mutations have been identified in the *PMS1* and *PMS2* genes (Nicolaides et al., 1994; Hamilton et al., 1995).

Genotype/phenotype correlation studies in HNPCC are still in a preliminary phase. Mutations in the *MSH2* and *MLH1* genes seem to be equally common in families selected through the Amsterdam criteria and together account for 50–80% of these families (Nystrom-Lahti et al., 1994; Liu et al., 1996). The empirical cumulative risk of developing colorectal cancer in putative gene carriers has been estimated as 80–90% (Mecklin et al., 1994). A risk analysis in 19 families with mutations at *MSH2* or *MLH1* showed no between-loci heterogeneity of age-specific risks for both colorectal and endometrial cancers (Vasen et al., 1996).

In the present study, we report the screening for germline mutations of the *MSH2*, *MLH1*, and *GTBP* genes in 14 unrelated Italian HNPCC patients and describe the phenotypic characteristics of these and two additional kindreds, whose mutations were characterized in previous studies (Wijnen et al., 1995, 1996).

Materials and methods

Family ascertainment. The 16 families included in this study were identified through a systematic investigation of the familial occurrence of colorectal cancer in patients hospitalized at the National Tumor Institute of Milan. Probands were asked, through a simple questionnaire, about the number and age of their relatives with large bowel or extracolonic tumors. Those with one or more affected first-degree relatives were interviewed for complete pedigree drawing and ascertainment of HNPCC status.

Isolation of DNA and RNA. Peripheral blood leukocytes lysed by Proteinase K digestion were used as source of genomic DNA. Total RNA was isolated from Ficoll-purified lymphocytes by extraction with a 14-M solution of guanidine salts and urea and RNA-binding resin purification.

RT-PCR, IVSP assay, and SSCP analysis. cDNAs were generated from RNA by reverse transcription (RT) with random hexamers and used as templates for polymerase chain reaction (PCR) amplifications with *MSH2* and *GTBP* primers. RT-PCR products were analyzed by in vitro synthesized protein (IVSP) assay as described by Powell et al. (1993) by using the TNT Coupled Reticulocytes Lysate Systems reaction kit (Promega, Madison, WI).

For *MSH2*, single-strand conformation polymorphism (SSCP) analysis was performed on cDNA samples amplified by two sequential rounds of PCR. Labeling of PCR products and electrophoretic conditions were as previously described (Donghi et al., 1993). For *MLH1*, SSCP analysis was performed on DNA samples. PCR products were loaded onto 20% homogeneous Phast-gel (Pharmacia Biotech) and run on a Phast Systemy apparatus.
Gels were subsequently stained with a Phast-Gely Silver Kit (Pharmacia Biotech).

Primer sequences and PCR conditions for *MSH2* and *GTBP* analyses are available upon request. For *MLH1* screening, the primers reported by Han et al. (1995) were used.

Sequence analysis. PCR products were fractionated on agarose gels and purified by extraction with phenol chloroform. Sequencing reactions were carried out by the dideoxy-mediated chain termination method by using ^{32}P-end-labeled primers and analyzed by electrophoresis on 6% polyacrylamide gels and autoradiography. Each mutation was verified by two independent PCR and sequencing reactions.

Statistical analysis. Age-of-onset comparison between groups was done with the Mann-Whitney U test.

Results

Mutation characterization. To screen the *MSH2* gene, we used different strategies, including IVSP assay, SSCP analysis, and direct sequence analysis. Lymphocyte RNA was used as a template to generate by RT-PCR three overlapping fragments, each of approximately 1 kb, spanning the entire coding sequence, and the PCR products were transcribed and translated in vitro. RT-PCR products of shorter length (200–400 bp) were analyzed by SSCP. Fragments exhibiting altered migration patterns were sequenced to determine the nature of the observed variations. Even when no anomalous electrophoretic patterns were observed, the sequence of the entire coding region was determined by using the RT-PCR products as templates.

In the proband of family 357, a shortened polypeptide was detected by IVSP assay of the 3'-end cDNA fragment (Fig. 1A). In addition, this individual also displayed altered SSCP patterns for two overlapping fragments in the same region (data not shown). Sequence analysis identified one base

deletion at nucleotide 2294 (codon 765) in exon 14 (Fig. 1B). This mutation caused a frame shift, leading to a translation stop signal at codon 811.

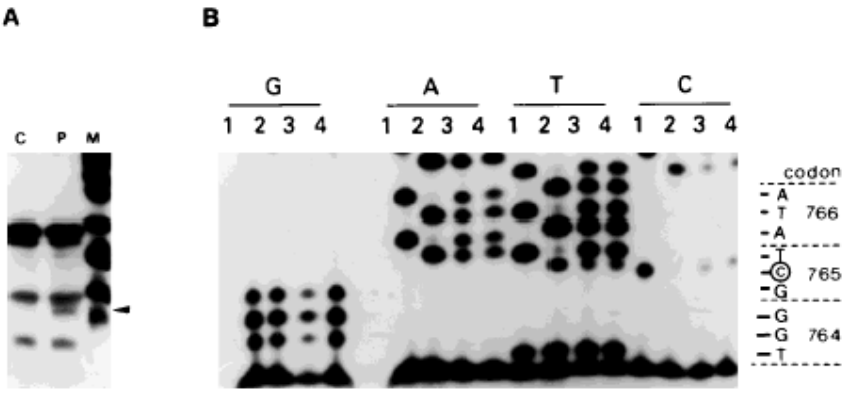

Figure 1. **A**: IVSP analysis of MSH2 cDNA fragment spanning the 3' end of coding region (nt. 2075-2850, according to Leach et al. 1993) in the proband (P) of family 357. The arrowhead indicates the shortened polypeptide product due to the presence of a truncating mutation. C, wild-type control. **B**; Sequence analysis of MSH2 mutation in family 357. The following templates were used: 1, wild-type DNA; 2, anomalous band excised from SSCP gel of proband DNA (see Fig. 3); 3 and 4, genomic DNA from two affected individuals from family 357. The deleted base in codon 765 is indicated in a circle.

In two cases, from families 349 and 667, a shortened RT-PCR fragment spanning exon 5 was observed. Sequence analysis of intron–exon junctions in genomic DNA identified the previously reported A-to-T transversion at the splice donor site of exon 5 (Liu et al., 1994, 1996). These findings were confirmed by using the restriction digest assay described by Froggatt et al. (1995) (data not shown).

To search for mutations in the *MLH1* gene, each of the 19 exons was amplified from genomic DNA and analyzed by no radioactive SSCP. Three base changes were identified: two introducing missense mutations and one affecting an RNA splicing site. In the proband of family 338, a nucleotide change from GCA to CCA at codon 128 in exon 5 caused an amino-acid change from alanine to proline. In the proband of family 311, a substitution from GGT to GAT at codon 244 in exon 9 caused an amino-acid change from glycine to aspartic acid. In the proband of family 668, an A-to-G transition in the consensus splice donor sequence of intron 6 was detected. RT-PCR of the corresponding cDNA encompassing *MLH1* exon 6 identified a shortened fragment (Fig. 2). Sequencing of the altered fragment revealed that all but the first three nucleotides of exon 6 were lost (data not shown), probably due to the presence of a cryptic splice donor site (GTGGAG) immediately downstream from the first three bases in *MLH1* exon 6.

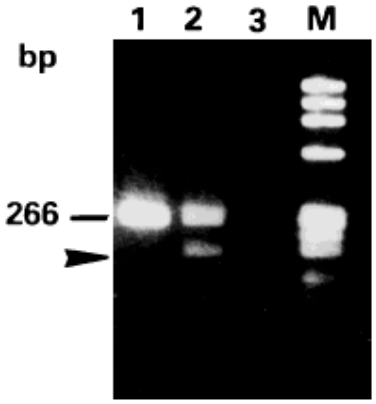

Figure 2. RT-PCR encompassing MLH1 exon 6. Lymphocyte cDNAs from a control (lane 1) and from the proband of family 668 (lane 2) were PCR amplified by using primers that detect a 266-bp fragment spanning from codon 139 in exon 5 to codon 228 in exon 9. The anomalous fragment detected in the cDNA of the family 668 proband is indicated by an arrowhead. Lane 3: Control reaction with no cDNA added.

The presence of the *MLH1* nucleotide changes was verified by SSCP analysis in 40 unrelated non-HNPCC individuals. None of these subjects displayed anomalous electrophoretic patterns.
The six *MSH2* and *MLH1* characterized mutations are summarized in Table 1.

TABLE 1. Mutations of MMR Genes in Italian HNPCC Families

	Family	Exon affected	Codons affected	Nucleotide change	Predicted protein change
MSH2	357	14	765	2294delC	Frameshift to stop cod 811
	349	5	265–314	GTA → GTT at splice donor site	In-frame deletion of exon 5
	667	5	265–314	GTA → GTT at splice donor site	In-frame deletion of exon 5
MLH1	338	5	128	GCA → CCA	Alanine to proline
	311	9	244	GGT → GAT	Glycine to aspartic acid
	668	6	150–182	GTA → GTG at splice donor site	Sub-total deletion of exon 6 (stop at cod 191)

Two previously described common polymorphisms in the *MLH1* gene (Liu et al., 1995; Tannergard et al., 1995) were also detected. One caused an amino-acid substitution in exon 8, changing Ile219 to Val (ATC to GTC; observed frequencies: A = 0.75/G = 0.25). The other consisted of an A-to-G substitution at position 219 in intron 14 (observed frequencies: a =0.42/g = 0.58).

The screening of the *GTBP* gene was accomplished by using the IVSP technique, as described for *MSH2*. Four overlapping RT-PCR fragments were examined. In none of the examined cases were truncated polypeptides observed.

Family analysis. We considered 16 families, including the 14 families screened for mutations at the *MSH2*, *MLH1*, and *GTBP* genes and two other families with previously characterized mutations (Wijnen et al., 1995, 1996).

In each family, we selected the three colorectal cancer cases that had been first diagnosed and on which the ascertainment of HNPCC status was based; these cases were referred to as ‘‘index cases’’ and the generations in which they occurred, ‘‘index generations.’’ The ages at tumor onset (or age at death) in the index generations of each family are given in Table 2.

Four families had a mutation in the *MSH2* gene, including large kindred (family 667) with three founders (two brothers and a cousin), each of whom had seven children and many grandchildren (19 cases of colorectal cancer). One family (219) was a ‘‘non-Amsterdam’’ case because the transmitting parent had a cancer of the uterus instead of a colon cancer. In these families, 25 individuals were screened for the identified mutations, mostly from family 357 (Fig. 3). The youngest affected mutation carriers were a brother and a sister, both of whom developed cancer of the colon at 28 years of age. The oldest unaffected mutation carrier was a woman of 43 years (endoscopically controlled) from family 357. All tested patients with colon cancer had the identified pathogenetic mutations.

TABLE 2. Age of Onset (or Age of Death) of Colorectal and Extracolonic Cancers in Italian HNPCC Families

Family	Mutated gene[a]	Index cases[b]	Other cases[c]	Remarks
219	MSH2	28,* 28, 31	co 54, ut [56, 59], st 38	
349	MSH2	30, 47, 61	bl 42	
357	MSH2	32, 40, 54	pa 31, ic 33	both founders with colon cancer at 54 and 61.
667	MSH2	38,* 40, 47	co [39, 40, 51, 61, 63], ut 36	two other branches with 11 cancer of the colon and 5 of the uterus.
202	MLH1	37, 42, 45*	co 49,* ut 52	
311	MLH1	36, 68,* 77	ki 64	founders were first cousins and both had colon cancers, at 77 and 93.
338	MLH1	32,* 36, 54	co [34, 35]	
668	MLH1	24, 45, 60*	co 58	
296	—	46, 66, 69	co [63, 66], sk 60	a cousin of the proband with colon cancer at 30.
298	—	33, 52, 55	co [54, 55, 61,* 62]	founder dying from cancer of the liver, at 61.
308	—	47, 68, 89	co [85, 91*], st [74, 78]	six other descendants with cancer, including a prostate carcinoma at 33 and a case with multiple neoplasias.
313	—	43,* 46, 53	br 58	
319	—	39, 49, 72	co 46, ut [59, 59], ic 37	both founders with cancer: breast at 65 and colon at 72.
327	—	42,* st 63, 81	ky 65	
350	—	26, 35,* 41	ut [46, 60]	second branch with two colon cancers, at 45 and 46, and a cancer of the uterus (unknown age).
353	—	42, 60, 71	—	two cousins of the proband with cancer of the colon at 22 and 52.

[a]All mutations identified in the present study except for families 219 (Wijnen et al.,1995) and 202 (Wijnen et al., 1996).
[b]All with colorectal cancer unless otherwise specified (st, stomach); a star indicates the occurrence of multiple primary malignancies, with the age of first cancer reported.
[c]Additional cancer cases in the two index generations: co, colorectal; ut, uterus; st, stomach; bl, bladder; pa, pancreas; ic, intracranial; ki, kidney; sk, skin; br, breast.

Four other families, all fulfilling the Amsterdam criteria, had a mutation in the *MLH1* gene. A total of 18 relatives were screened for the identified mutations. The youngest patient with a confirmed mutation developed colon cancer at the age of 32 years. The oldest unaffected mutation carrier was

a woman, healthy at 50 years. In these families, all tested patients with colon cancer had the identified pathogenetic mutation.

All but one of the eight families that were negative for all three mismatch repair loci tested fulfilled the Amsterdam criteria.

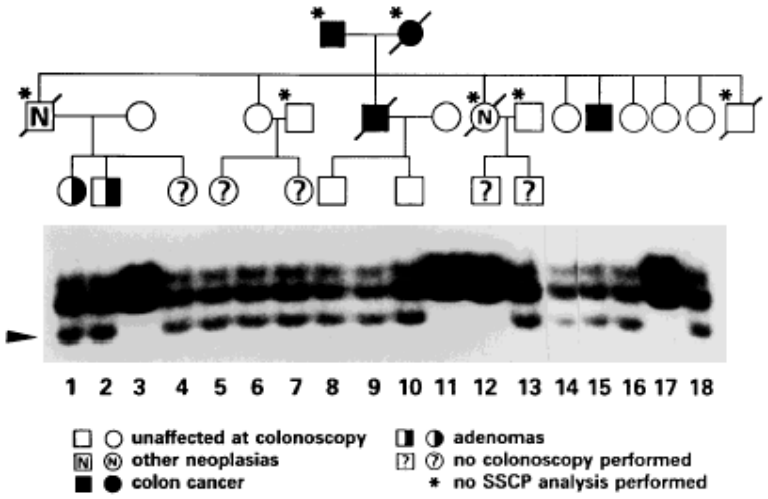

Figure 3. Molecular diagnosis in HNPCC family 357. Lymphocyte cDNA from affected and healthy individuals were amplified by PCR with primers flanking the mutation site in MSH2 exon 14 and subjected to SSCP analysis as described in the text. The arrowhead indicates the anomalous band in individuals carrying the mutation. The two individuals with extracolonic neoplasia died of cancer of the brain and pancreas.

The overall mean age of onset of colorectal cancer (n = 67) in the two index generations of the 16 families was 50.6 ± 15.7 years. The mean age of onset of cancers of the uterus (n = 8, two with mutation confirmed) was 53.4 ± 8.5 years. Other cancers were as follows: stomach, four; kidney, two (one with mutation confirmed); intracranial, two (one with mutation inferred); bladder, one; pancreas, one (mutation inferred); skin, one; and breast, one.

When we compared the means of the age of tumor onset in the index cases, we found a significantly lower value in the families with mutation than in the families without mutation (43.0 ± 13.5 vs. 53.7 ± 15.9 years, $P = 0.014$). This difference was further investigated by considering the number of index cases with age of tumor onset <50 years in each family. In the group with an identified mutation, three families had all three index cases <50 years, four families had two index cases <50 years, and one family had a single index case <50 years. The corresponding figures in the group without mutation were as follows: one family with all three index cases with age of tumor onset <50 years, two families with two index cases <50 years, and five families with a single index case <50 years.

Discussion

The isolation and characterization of the human MMR genes have enabled the establishment of mutation-detection strategies necessary to facilitate presymptomatic diagnosis of HNPCC. This task is, however, complicated by the presence of pathogenic mutations in several genes and by the heterogeneity of the associated lesions. The sensitivity of mutation detection is therefore greatly increased by the simultaneous use of multiple screening approaches (Liu et al., 1996; Nystrom-Lahti et al., 1996).

We started our mutational screening of the *MSH2* gene in HNPCC families with three different methods: IVSP assay, SSCP analysis, and direct sequencing. Subsequently, two additional MMR genes, *MLH1* and *GTBP*, were examined by using single-method strategies. SSCP analysis was chosen to screen *MLH1* because of the relatively frequent finding in this gene of mutations causing amino acid substitutions in addition to truncating ones (Han et al., 1995; Tannergard et al., 1995; Liu et al., 1996; Nystrom-Lahti et al., 1996; Wijnen et al., 1996). Conversely, the IVSP assay, which is a relatively simple method to identify truncating mutations in large genes, was selected as a primary screening test for *GTBP*. In fact, most *GTBP* alterations reported to date in hypermutable cell lines are frame shifts or nonsense mutations (Papadopoulos et al., 1995).

Among the 14 Italian HNPCC families screened at the *MSH2*, *MLH1*, and *GTBP* loci, we identified three mutations in the *MSH2* gene and three in the *MLH1* gene. Of these, one was a frame-shift mutation and three led to total or ‘‘subtotal’’ exon deletions. These mutations are clearly inactivating; they are predicted to lead to the complete loss of function of the protein encoded by the altered allele.

The remaining two mutations identified in *MLH1* were missense mutations. In kindred 338, an Ala128→Pro causes the substitution of a no hydrophobic amino acid, unfavorable to a helix, for a small hydrophobic amino acid, which is favorable to a helix. In kindred 311, a Gly344→Asp causes the substitution of a polar (acid) amino acid for a small no polar one. At present, no functional assay systems to determine the biological consequence of missense mutations in this gene are available, and thus the relationship between missense mutations and genetic predisposition to cancer is difficult to determine. However, although the observed substitutions affected residues that are not conserved between human and yeast (Bronner et al., 1994; Papadopoulos et al., 1994), they introduced substantial amino-acid changes and were not found in 40 unrelated non-HNPCC

individuals. This finding suggests that the two alterations are unlikely to represent simple polymorphisms.

In total, including the two families previously reported (Wijnen et al., 1995, 1996), mutations in MMR genes were identified in 50% (8 of 16) of the Italian HNPCC families. More precisely, both *MSH2* and *MLH1* were found to be associated with 25% of examined cases, i.e., in four families each. These frequencies are similar to those reported in other populations, ranging from 14% to 31% for *MSH2* (Wijnen et al., 1995; Liu et al., 1996; Lu et al., 1996; Moslein et al., 1996) and from 20% to 35% for *MLH1* (Han et al., 1995; Tannergard et al., 1995; Liu et al., 1996; Moslein et al., 1996; Wijnen et al., 1996). The exceptionally high rate of *MLH1* mutation detection (90%) reported in Finnish families by Nystrom-Lahti et al. (1996) was biased by the presence of founder effects in the examined population. As to *MLH1*, the results obtained are in keeping with linkage analyses estimating that approximately 30% of HNPCC families are linked to this gene (Bronner et al., 1994; Nystrom-Lahti et al., 1994). In contrast, the proportion of families carrying germline *MSH2* mutations was lower than the 50–60% predicted by linkage studies (Liu et al., 1994; Nystrom-Lahti et al., 1994).

Such a discrepancy might be due to the occurrence, in a fraction of the examined families, of mutations that are not detectable by the screening techniques adopted in this and other studies, for example, large intragenic deletions or rearrangements. Alternatively, in some chromosome arm-2p-linked HNPCC families, the actual mutations may affect a gene syntenic to *MSH2*. An obvious candidate is *GTBP*, which maps closely to *MSH2* (Drummond et al., 1995; Palombo et al., 1995). However, we were unable to identify *GTBP* germline mutations by IVSP assay screening of our families. Although it is not possible to rule out the presence of other types of mutations not detectable by this technique, our results are consistent with those of Liu et al. (1996), who found no *GTBP* mutations in a group of 48 HNPCC families screened by both the IVSP assay and direct sequence analysis.

Whereas the number of families with *MSH2* or *MLH1* mutations investigated in the present study is too small to assess possible differences in phenotypic expression among loci or mutations, an unexpected association concerning age of onset was observed. There was a significantly earlier mean age of onset of colorectal cancers (about 10 years) in the index cases of the families with either a *MSH2* or a *MLH1* mutation compared with the index cases of families negative at both loci. In addition, seven of eight families (87%) of the first group (with mutation) had two or three index

cases with age of tumor onset <50 years, whereas the corresponding figure in the second group was three of eight (37%). Notably, the mean age of tumor onset in families with identified mutation (43 years) corresponded to other published values (Lynch and Lynch, 1995).

On the basis of these observations, we suggest that mutations of *MSH2* and *MLH1* cause a more severe phenotype with early age of tumor onset, whereas mutations of other MMR genes still to be identified (or in other genes conferring susceptibility to colorectal cancer) are associated with a significantly higher mean age of tumor onset. Alternatively, some *MSH2* or *MLH1* gene mutations that confer a milder phenotype might not be detected by the screening methods used in the present study. We are analyzing the implications of these findings in the definition of screening protocols for mutation analysis in HNPCC families.

Acknowledgments

We thank Maria T. Radice for technical assistance, Mr. Mario Azzini for preparation of the figures, and Mrs. Anna Grassi and Miss Cristina Mazzadi for secretarial assistance. We are grateful to Prof. Franco Rilke for critical reading of the manuscript.
This work was supported in part by funds from the Italian Ministry of Health and the Italian Association and Foundation for Cancer Research (A.I.R.C./F.I.R.C.). G.B. is supported by a fellowship from A.I.R.C.

References

Aaltonen LA, Sankila R, Mecklin J-P, Jarvinen H, Pukkala E,Peltomaki P, de la Chapelle A (1994) A novel approach to estimate the proportion of hereditary non-polyposis colorectal cancer of total colorectal cancer burden. Cancer Detect Prev 18:57–63.

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, Kane M, Earabino C, Lipford J, Lindblom A, Tannergard P, Bollag RJ, Godwin AR, Ward DC, Nordensjold M, Fishel R, Kolodner R, Liskay RM (1994) Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. Nature 368:258–261.

Buerstedde J-M, Alday P, Torhorst J, Weber W, Muller H, Scott R (1995) Detection of new mutations in six out of 10 Swiss families by genomic sequencing of the hMSH2 and hMLH1 genes. J Med Genet 32:909–912.

Donghi R, Longoni A, Pilotti S, Michieli P, Della Porta G, Pierotti MA (1993) Gene p53 mutations are restricted to poorly differentiated and undifferentiated carcinomas of the thyroid gland. J Clin Invest 91:1753–1760.

Drummond JT, Li G-M, Longley MJ, Modrich P (1995) Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science 268:1909–1912.

Fishel R, Lescoe MK, Rao MRS, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R (1993) The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. Cell 75:1027–1038.

Froggatt NJ, Joyce JA, Davies R, Evans DGR, Ponder BAJ, Barton DE, Maher ER (1995) A frequent *hMSH2* mutation in hereditary non-polyposis colon cancer syndrome. Lancet 345:727.

Hamilton SR, Liu B, Parsons RE, Papadopoulos N, Jen J, Powell SM, Krush AJ, Berk T, Cohen Z, Tetu B, Burger PC, Wood PA, Taqi F, Booker SV, Petersen GM, Offerhaus GJA, Tersmette AC, Giardiello FM, Vogelstein B, Kinzler KW (1995) The molecular basis of Turcot's syndrome. N Engl J Med 332:839–847.

Han H-J, Maruyama M, Baba S, Park J-G, Nakamura Y (1995) Genomic structure of human mismatch repair gene, hMLH1, and its mutation analysis in patients with hereditary non-polyposis colorectal cancer (HNPCC). Hum Mol Genet 4:237–242.

Kee F, Collins BJ (1992) Families at risk of colorectal cancer: Who are they? Gut 33:787–790.

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Peltomaki P, Sistonen P, Aaltonen LA, Nystrom-Lahti M, Guan X-Y, Zhang J, Meltzer PS, Yu J-W, Kao F-T, Chen DJ, Cerosaletti KM, Fournier REK, Todd S, Lewis T, Leach RJ, Naylor SL, Weissenbach J, Mecklin J-P, Jarvinen H, Petersen GM, Hamilton SR, Green J, Jass J,Watson P, Lynch HT, Trent JM, de la Chapelle A, Kinzler KW, Vogelstein B (1993) Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. Cell 75:1215–1225.

Liu B, Parsons RE, Hamilton SR, Petersen GM, Lynch HT, Watson P, Markowitz S, Willson JKV, Green J, de la Chapelle A, Kinzler KW, Vogelstein B (1994) hMSH2 mutations in hereditary nonpolyposis colorectal cancer kindreds. Cancer Res 54:4590–4594.

Liu B, Nicolaides NC, Markowitz S,Willson JKV, Parsons RE, Jen J, Papadopoulos N, Peltomaki P, de la Chapelle A, Hamilton SR, Kinzler KW, Vogelstein B (1995) Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. Nature Genet 9:48–55.

Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Peltoma¨ki P, de la Chapelle A, Hamilton SR, Vogelstein B, Kinzler KW (1996) Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nature Med 2:169–174.

Lu AL, Clark S, Modrich P (1993). Methyl-directed repair of DNA base-pair mismatches in vitro. Proc Natl Acad Sci USA 80:4639– 4643.

Lu SL, Akiyama Y, Nagasaki H, Nomizu T, Ikeda E, Baba S, Ushio K, Iwama T, Maruyama K, Yuasa Y (1996) Loss or somatic mutations of *hMSH2* occur in hereditary nonpolyposis colorectal cancers with *hMSH2* germline mutations. Jpn J Cancer Res 87:279–287.

Lynch HT, Lynch JF (1995) Clinical implications of advances in the molecular genetics of colorectal cancer. Tumori 81:19–29.

Lynch HT, Smyrch TC, Watson P, Lanspa SJ, Lynch JF, Lynch PM, Cavalieri RJ, Boland CR (1993) Genetics, natural history, tumor spectrum, and pathology of hereditary non-polyposis colorectal cancer: An update review. Gastroenterology 104:1535–1549.

Maliaka YK, Chudina AP, Belev NF, Alday P, Bochkov NP, Buerstedde JM (1996) CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for HNPCC mutations. Hum Genet 97:251– 255.

Mecklin J-P, Svendsen LB, Peltomaki P, Vasen HFA (1994) Hereditary nonpolyposis colorectal cancer. Scan J Gastroenterol 29:673– 677.

Moslein G, Tester DG, Lindor NM, Honchel R, Cunningham JM, French AJ, Halling KC, Schwab M, Goretzki P, Thibodeau SN (1996) Microsatellite instability and mutation analysis of hMSH2 and hMLH1 in patients with sporadic, familial and hereditary colorectal cancer. Hum Mol Genet 5:1245–1252.

Nicolaides NC, Papadopoulos N, Liu B, Wei Y-F, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD, Venter JC, Dunlop MG, Hamilton SR, Petersen GM, de la Chapelle A, Vogelstein B, Kinzler KW (1994) Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 371:75–80.

Nystrom-Lahti M, Parsons R, Sistonen P, Pylkkanen L, Aaltonen LA, Leach FS, Hamilton SR, Watson P, Bronson E, Fusaro R, Cavalieri J, Lynch J, Lanspa S, Smyrk T, Lynch P, Drouhard T, Kinzler KW, Vogelstein B, Lynch HT, de la Chapelle A, Peltomaki P (1994) Mismatch repair genes on chromosomes 2p and 3p account for a major share of hereditary nonpolyposis colorectal cancer families evaluable by linkage. Am J Hum Genet 55:665.

Nystrom-Lahti M, Wu Y, Moisio AL, Hofstra RMW, Osinga J, Mecklin JP, Jarvinen HJ, Leisti J, Buys CHCM, de la Chapelle A, Peltomaki P (1996) DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum Mol Genet 5:763–769.

Palombo F, Gallinari P, Iaccaraino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ, Jiricny J (1995) GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science 268:1912–1914.

Papadopoulos N, Nicolaides NC, Wei Y-F, Ruben SM, Carter KC, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD, Venter JC, Hamilton SR, Petersen GM,Watson P, Lynch HT, Peltomaki P, Mecklin J-P, de la Chapelle A, Kinzler KW, Vogelstein B (1994) Mutation of a *mutL* homolog in hereditary colon cancer. Science 263:1625–1629.

Papadopoulos N, Nicolaides NC, Liu B, Parsons R, Lengauer C, Palombo F, D'Arrigo A, Markowitz S, Willson JKV, Kinzler KW, unstable cells. Science 268:1915–1917.

Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, Hamilton SR, Vogelstein B, Kinzler KW (1993) Molecular diagnosis of familial adenomatous polyposis. N Engl J Med 329:1982– 1987.

Strand M, Prolla TA, Liskay RM, Petes TD (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365:274–276.

Tannergard P, Lipford JR, Kolodner R, Frodin JE, Nordenskjold M, Lindblom A (1995) Mutation screening in the hMLH1 gene in Swedish hereditary nonpolyposis colon cancer families. Cancer Res 55:6092–6096.

Vasen HFA, Mecklin J-P, Khan PM, Lynch HT (1991) Hereditary nonpolyposis colorectal cancer. Lancet 338:877

Vasen HFA, Wijnen J, Meera Khan P, Menko FH, Kleibeuker JK, Taal BG, Griffioen G, Nagengast FM, Meijers-Heijboer H, Bertario L, Varesco L, Bisgaard M, Mohr J, Fodde R (1996) Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. Gastroenterology 110:1020–1027.

Wijnen J, Vasen H, Khan PM, Menko FH, Van der Klift H, Van Leeuwen C, Van den Broek M, Van Leeuwen-Cornelisse I, Nagengast F, Mejiers-Heijbor A, Lindhout D, Griffionen G, Cats A, Kleibeuker J, Varesco L, Bertario L, Bisgaard ML, Mohr J, Fodde R (1995) Seven new mutations in hMSH2, an HNPCC gene, identified by denaturing gradient gel electrophoresis. Am J Hum Genet 56:1060–1066.

Wijnen J, Khan PM, Vasen H, Menko FH, Van der Klift H, Van den Broek M, Van Leeuwen-Cornelisse I, Nagengast F, Meijers- Heijboer A, Lindhout D, Griffionen G, Cats A, Kleeibeuker J, Varesco L, Bertario L, Bisgaard ML, Mohr J, Kolodner R, Fodde R (1996) Majority of hMLH1 mutations responsible for hereditary nonpolyposis colorectal cancer (HNPCC) cluster at the exonic region 15–16. Am J Hum Genet 58:300–307.

Chapter 5

A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor

Gazzoli I, Loda M, Garber J, Syngal S and Kolodner RD.

Cancer Research 2002 Jul 15; 62(14):3925-8.

Abstract

Fourteen suspected hereditary nonpolyposis colorectal carcinoma cases with microsatellite unstable(microsatellite instability-high; MSI-H) tumors but no germ-line MSH2, MSH6, or MLH1 mutations were examined for hypermethylation of CpG sites in the critical promoter region of MLH1. The methylation patterns were determined using methylation-specific PCR and by sequence analysis of sodium bisulfite-treated genomic DNA. In one case, DNA hypermethylation of one allele was detected in DNA isolated from blood. In the MSI-H tumor from this case, the unmethylated MLH1 allele was eliminated by loss of heterozygosity, and the methylated allele was retained. This biallelic inactivation resulted in loss of expression of MLH1 in the tumor as confirmed by immunohistochemistry. These results suggest a novel mode of germ-line inactivation of a cancer susceptibility gene.

Introduction

Approximately 60–90% of CpG dinucleotides are methylated on cytosine in the human genome, although unmethylated GC-rich regions are frequently associated with transcriptionally active genes. Epigenetic alteration of the human genome can affect cytosine methylation and chromatin structure [1, 2, 3]. A number of studies have demonstrated that DNA methylation can directly prevent the binding of transcription factors, such as E2F, cyclic AMP-responsive element binding protein, and USF [4, 5, 6], or act through the binding of methylated CpG-binding proteins to induce chromatin configurations that interfere with the transcription machinery [1, 7]. As a consequence, changes in the methylation of CpG-rich regions can result in altered gene expression that is independent of genetic alteration of either coding or regulatory sequences. In the last decade, a number of studies have reported that aberrantly hypermethylated CpG islands are associated with transcriptional silencing of tumor suppressor genes in sporadic cancers (reviewed in Refs. [2, 3]). Thus, methylation-associated transcriptional silencing is an alternative to mutational inactivation as a cause of loss of tumor suppressor gene function. Consequently, the analysis of DNA cytosine methylation patterns provides an alternative strategy for the study of tumor suppressor gene inactivation in cases where mutations are not observed. One of many such examples is the MMR[3] gene MLH1, where hypermethylation of the MLH1 promoter region is associated with loss of expression and appears to underlie the majority of MMR-defective sporadic cancer cases [8, 9, 10, 11].

In a previous study [12, 13], we analyzed 70 HNPCC cases (which met Amsterdam, Modified Amsterdam, HNPCC-like, or Bethesda criteria) for germ-line defects in MSH2, MLH1, and MSH6 genes. For 48 of the cases, a tumor sample was available, and the MSI status of the tumor was determined; 14 cases were MSI-H and showed a germ-line mutation in MSH2 or MLH1, and 14 cases were MSI-H but did not have a germ-line mutation in MSH2, MSH6, or MLH1 [14]. The observation of HNPCC cases that have MSI-H tumors but no germ-line mutation in a known MMR gene, as detected by DNA sequencing or other methods for identifying point mutations, has been reported by many studies. Some of these types of case have been attributed to germ-line deletion mutations in MSH2 or MLH1 as well as uncharacterized mutations that cause loss of expression of MLH1 [15, 16]. In the present study, the mutation-negative 14 HNPCC cases were examined for hypermethylation of CpG sites in the critical promoter region of MLH1 (-290 to -180 bp relative to the ATG), which shows a strong correlation with the loss of MLH1 expression [17]. DNA from

blood of one of these 14 cases was found to have one methylated MLH1 allele, and analysis of the MSI-H tumor from this case demonstrated that LOH eliminated the unmethylated MLH1 allele, resulting in loss of expression of MLH1.

Materials and Methods

Tissue and DNA samples. The HNPCC cases used in the studies described here have been described in detail elsewhere (12 , 13). DNA from blood was isolated as described previously and analyzed for the presence of mutations in MSH2, MLH1, and MSH6 by DNA sequencing as reported previously. Tumor DNA from each case was analyzed for MSI using the National Cancer Institute recommended five microsatellite markers as described previously [14]. The 14 cases studied here comprised all 14 of the MSI-H cases that did not have mutation in MSH2, MLH1, or MSH6 described previously [14]. Control genomic DNA in which MLH1 was not methylated was from blood of the case DF000268 containing the germ-line MSH2 IVS5 + 3A>T splice site mutation [12]. Control genomic DNA in which MLH1 was fully methylated was isolated from SW48 cells (American Type Culture Collection CCL-231) with a Puregene DNA isolation kit (Gentra Systems).

Analysis of MLH1 methylation using MS-PCR. One μg of total genomic DNA was modified by sodium bisulfite treatment using a CpGenome Modification kit (Intergen) according to the manufacturer's instructions and suspended in 10 μl of 1x TE buffer [10 mM Tris-HCl (pH 7.6), 1 mM EDTA]. The methylation status of the MLH1 gene was then determined by MS-PCR using primers specific for both methylated and unmethylated sequences as essentially described previously [18]. PCR reactions (20 μl) contained 1.5 μl of modified DNA and were performed using an Advantage GC2 PCR kit (Clontech). The thermocycling conditions used were 1 cycle of 94°C for 2 min; 3 cycles of 94°C for 30 s, 60°C for 30 s, 68°C for 30 s; 36 cycles 94°C for 30 s, 53°C for 30 s, 68°C for 30 s, and 1 cycle of 68°C for 2 min. Three μl of each PCR reaction were then analyzed by electrophoresis through nondenaturing 10% polyacrylamide gels, which were stained with ethidium bromide and visualized under UV illumination.

Analysis of MLH1 by DNA sequencing. To analyze the methylation status of MLH1 by direct DNA sequencing, genomic DNA was modified by treatment with sodium bisulfite as described above. The promoter region of the MLH1 gene was amplified using two rounds of PCR with nested

primers using an Advantage GC2 PCR kit. The first round of amplification used the primers 5-GTTTGAGAAGTGTTAAGTATTTTTTT-3' and 5'-CAAATAACCCCTACCACAAAC-3', and the second round of amplification used the primers 5'-GTATTTTTTTTGTTTTGYGTTAG-3'and 5'-CTATTAATTAAACAACTTAAATACCAATC-3'. The final PCR product then was treated with shrimp alkaline phosphatase and exonuclease I (Amersham) to remove excess PCR primers and nucleotides, and the resulting PCR product was sequenced using a PE/ABI 377 sequencer and dye terminator chemistry (PE Applied Biosystems).

To analyze the presence of a polymorphism in the 5'-untranslated region of the MLH1 gene, genomic DNA from tumor and blood samples was used as template in PCR reactions using primers 5'-GAAAACTAGAGCCTCGTCGACTT-3' (sense) and 5'-TAGCATTAGCTGGCCGCTGGATAAC-3' (antisense) essentially as described previously. The resulting PCR product was then purified and sequenced using a PE/ABI 377 sequencer and dye terminator chemistry (PE Applied Biosystems), and the sequence data obtained was analyzed using Sequencher 3.1 software (GeneCodes, Inc.).

Immunohistochemistry (IHC) analysis of MLH1 expression. Tissue sections were stained with anti-MLH1 antibody (G168-728; PharMingen) as previously described [19]. The results of immunohistochemistry analysis were kept blinded until the MLH1 methylation analysis was completed.

Results

In a previous study, we identified 14 suspected HNPCC cases (which met Amsterdam, Modified Amsterdam, HNPCC-like, or Bethesda criteria) that had MSI-H tumors but did not have a germ-line mutation in MSH2, MSH6, or MLH1. To understand the genetic basis for the MMR defect in these 14 HNPCC cases, we screened them for hypermethylation of the MLH1 gene. The MLH1 DNA methylation pattern was determined by sequencing PCR products amplified from bisulfite-modified DNA isolated from blood of all 14 mutation-negative cases and by MS-PCR using bisulfite-modified DNA from selected cases as template [18]. Cytosine methylation protects against bisulfite modification-induced conversion of cytosine to uracil. Consequently, during sequencing of modified DNA, unmethylated cytosines appear as thymines, methylated cytosines appear as cytosines, and mixtures appear heterozygous. Similarly, during MS-PCR, the primers specific for the sequences that result after bisulfite modification of methylated and unmethylated DNA generate

two fragments of different lengths containing the same region, 91 and 103 bp long, respectively. The size of the resulting fragments allows discrimination between methylated and unmethylated alleles by acrylamide gel electrophoresis. The region of MLH1 selected for analysis, -290 to -180 bp relative to the ATG, contains nine CpG sites close to the CCAAT box and different transcription factors binding sites, and hypermethylation of this region is associated with loss of expression of MLH1.

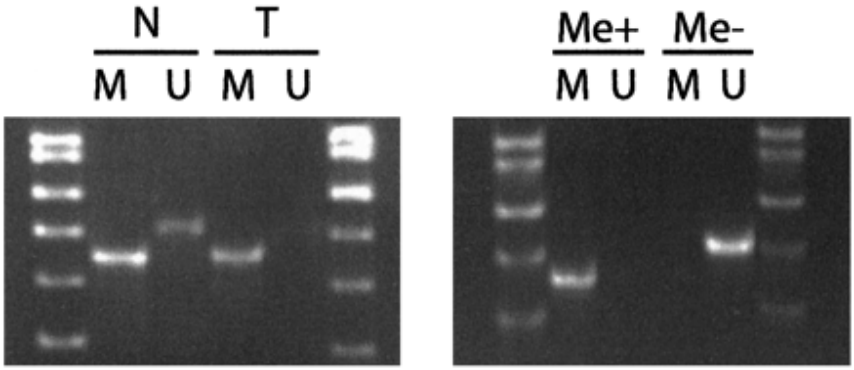

Fig. 1. Analysis of the methylation status of the *MLH1* gene in an HNPCC case using MS-PCR. *M* and *U*, the 91- and 103-bp PCR products amplified using primers specific for methylated and unmethylated alleles, respectively. *B*, blood DNA; *T*, tumor DNA; *Me+*, DNA from the SW48 cell line in which *MLH1* is fully methylated; *Me-*, DNA sample with a genetic alteration in *MSH2* gene, in which *MLH1* is unmethylated.

Using MS-PCR, one case (which met Bethesda criteria 4 alone; diagnosis at 25 years of age) was found to have both methylated and unmethylated MLH1 alleles when blood DNA was analyzed, and only methylated alleles when tumor DNA was analyzed (Fig. 1). Only unmethylated alleles were observed when unmethylated control DNA was analyzed, and conversely, only methylated alleles were observed when fully methylated DNA was analyzed from the SW48 cell line, which is MMR defective and does not express MLH1. To further analyze the hypermethylation status of this region, we PCR amplified the region from bisulfite-modified DNA and sequenced the resulting PCR products (Fig. 2). The chromatograms showed that the MLH1 promoter region was partially methylated (a mixture of C and T is seen at critical Cs in the region analyzed) in blood DNA and totally methylated (only C is seen at the critical Cs in the region analyzed) in tumor DNA; analysis of an area containing two CpG sites (8 and 15 bp downstream from the CAAT box located at position -282) is shown in Fig. 2. Sequencing of unmethylated and fully methylated control DNAs detected only unmethylated and methylated alleles, respectively (Fig. 2).

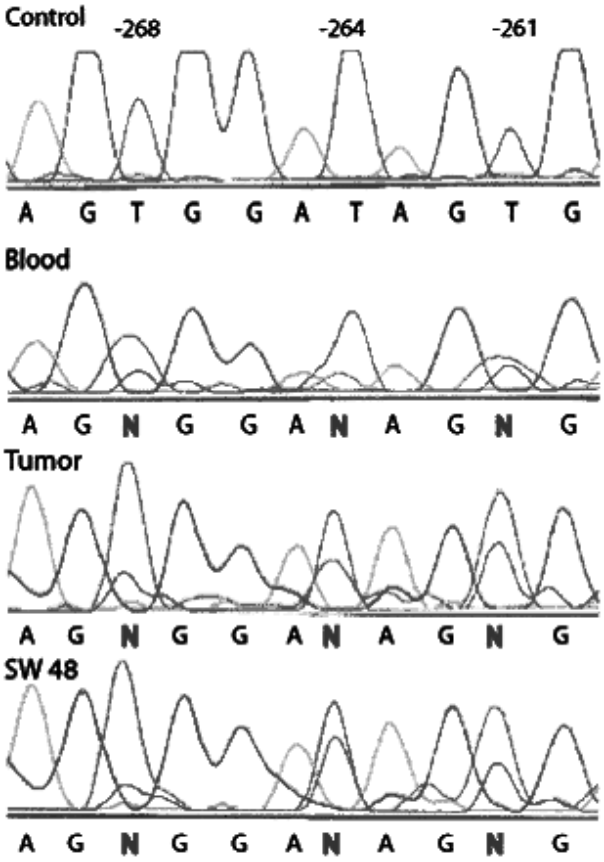

Figure 2. Sequence analysis of bisulfite-modified DNA. *Control*, unmethylated DNA from a *MSH2* mutant HNPCC case showing all Ts at positions -268, -264, and -261 attributable to complete modification of the DNA; *Blood*, DNA isolated from blood from the case showing a mixture of C and T (N) at positions -268, -264, and -261 attributable to partial modification of the DNA because of partial methylation of the DNA; *Tumor*, DNA isolated from tumor from the case showing an increase in the level of C (high peak) at positions -268, -264, and -261 attributable to reduced modification of the DNA because of increased methylation of the DNA; *SW48*, DNA from the SW48 cell line in which MLH1 is completely methylated showing a high level C at positions -268, -264, and -261 attributable to reduced modification because of complete methylation of the DNA.

Sequencing 20 cloned PCR products from each sample (data not shown) revealed a mixture of methylated and unmethylated clones from blood DNA and an increased proportion of methylated clones from tumor DNA, consistent with the results obtained by direct sequencing of PCR products shown in Fig. 2. The observation that ~50% of the DNA from blood from the mutation-negative HNPCC case described above showed MLH1 methylation by direct bisulfite sequencing suggests that the observed MLH1 methylation is unlikely to result from shedding of tumor cells into the blood [(18)], because this would likely yield lower levels on MLH1 methylation than observed here.

The above analysis documented an HNPCC case in which normal tissue contained a mixture of methylated and unmethylated MLH1 alleles, whereas tumor tissue contained only methylated MLH1 alleles. There are two possible explanations for the observation that the blood DNA contained one methylated and one unmethylated MLH1 allele while the tumor DNA only contained methylated alleles; the unmethylated allele could have been deleted or methylated in the tumor. Sequencing of the promoter region from unmodified DNA from this case revealed that the blood DNA was heterozygous for a previously described single nucleotide G→A polymorphism (Fig. 3;

Ref. [20]). When the same region of unmodified DNA from the tumor sample was sequenced, only the G allele was observed. This indicates that LOH had occurred in the tumor, resulting in the loss of the unmethylated allele (Fig. 3). Consistent with loss of the unmethylated, expressed MLH1 allele in the tumor, immunohistochemistry analysis showed that this case did not express MLH1 in the tumor (Fig. 4), whereas MSH2 was still expressed (data not shown). These observations are in contrast to the observation of biallelic methylation of MLH1 in sporadic colon tumors leading to loss of MLH1 expression [(21)].

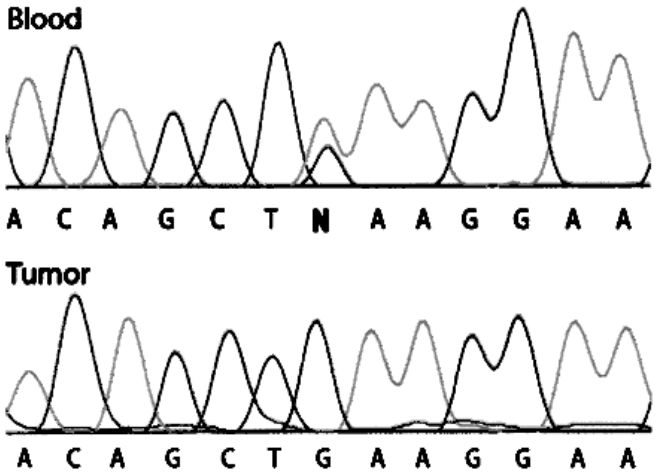

Fig. 3. Detection of LOH by sequence analysis of the 5' untranslated region of the *MLH1* gene. DNA isolated from blood was heterozygous (N) for a G→A polymorphism at position -93 bp, whereas DNA from the tumor only had a G at this position.

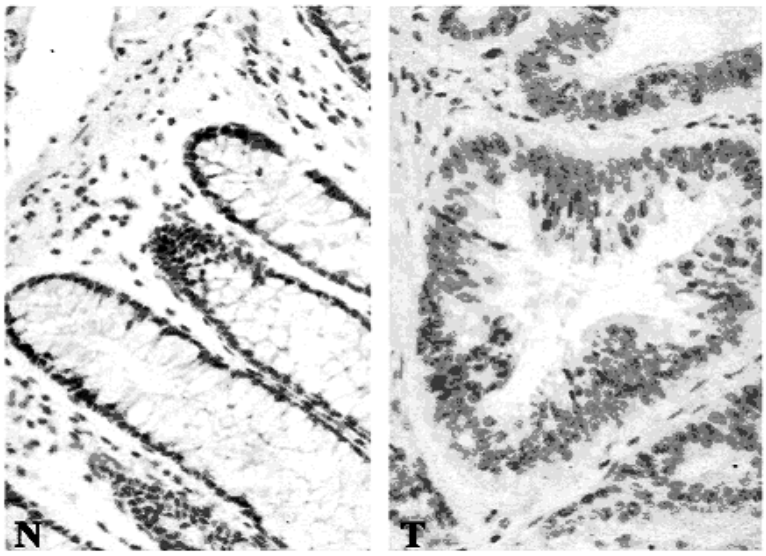

Fig. 4. Immunohistochemistry analysis of MLH1 expression. MLH1 staining is in *brown*. *N*, normal tissue; *T*, tumor tissue.

Discussion

A growing body of evidence has suggested that the majority of HNPCC families showing MSI can be accounted for by mutations in MSH2 and MLH1 (see Refs. [12, 13, 14, 16]). Initial studies documented missense, nonsense, frameshift, splice site, and specific deletion mutations in MSH2 and MLH1, and more recent studies have documented large MSH2 and MLH1 deletion mutations and uncharacterized mutations that eliminate expression of these genes [15, 16]. In the present study, we have identified an HNPCC case in which one allele of MLH1 was methylated in DNA isolated from normal tissue (blood) and somatic LOH of the unmethylated allele subsequently occurred, resulting in a tumor that did not express MLH1 and was MMR defective. These results raise the possibility that methylation and associated silencing of MLH1 could represent a germ-line alteration that underlies some HNPCC cases; however, samples from the parents were not available. Therefore, it was not possible to determine whether the methylated allele was actually inherited. We have detected one such case in our sample set of 14 mutation-negative, MSI-H HNPCC cases, indicating that germ-line silencing of MLH1 could underlie a small but significant proportion of HNPCC cases. Larger scale studies will be required to definitively identify other HNPCC cases similar to the one described here and to determine how frequently this unique mechanism of tumorigenesis occurs.

The region of MLH1 found to be methylated in our studies contains CpG sites close to the CCAAT box at position -282, and methylation of this region is associated with inactivation of the MLH1 gene. The CCAAT box is recognized by CBF, and binding of this factor effects MLH1 promoter activity. Functional studies have shown that methylation of this region interferes with binding of CBF, resulting in decreased transcription of MLH1 [22]. In addition to decreasing the binding of critical transcription factors, methylation of this region of MLH1 likely increases the affinity of methyl-binding proteins and histone deacetylases, leading to the assembly of inactive chromatin configurations. Consistent with this view, full reactivation in tumor cell lines requires treatment with both inhibitors of DNA methylation and histone deacetylation [23]. These observations support the view that the methylation of the region of MLH1 observed here is directly associated with or underlies the silencing of the gene observed in the HNPCC case reported here.

A number of examples of loss of function of different tumor suppressor genes attributable to somatic silencing of the genes associated with hypermethylation have been reported. The types of examples reported in sporadic cancer include biallelic methylation, somatic methylation associated

with LOH, and mutation associated with methylation (reviewed in Ref. [2]). These observations have led to the hypothesis that hypermethylation and associated silencing of tumor suppressor genes represents an alternate pathway of the two-hit tumor suppressor gene inactivation hypothesis [2, 3]. In those cases where methylation associated with either mutation or LOH has been observed, methylation appears to have been a somatic event. In at least one case, methylation of $p16^{INK4}$ was observed in both recurrent tumors and adjacent normal epithelium, suggesting that methylation in normal tissue might underlie subsequent tumor development [24]. In addition, age-related methylation of MLH1 has been observed in normal colonic mucosa associated with the development of colon tumors, with high-level methylation being associated with the development of MSI-H tumors, supporting the idea that the age-related development of MLH1 methylation is a precursor lesion [9]. Our results raise the possibility that methylation and associated silencing of a gene like MLH1 could actually represent a germ-line change that can underlie development of cancer. Alternately, somatic methylation could occur much more widely in normal tissues than thought previously [9]. This suggests a novel mechanism of tumorigenesis that nonetheless fits the two-hit tumor suppressor gene inactivation hypothesis, although larger scale studies are clearly required to verify our observations and determine how prevalent this mechanism is. In this regard, a recent study detected three cases of HNPCC where MLH1 expression was absent in the tumors, but a germ-line MLH1 mutation could not be found; however, MLH1 methylation was not analyzed in these cases [16].

Acknowledgments

This work was supported of a grant from the Starr Foundation and Grant CA06516 from the National Cancer Institute to the Dana-Farber Cancer Institute and Grant CA23100 from the University of California, San Diego Medical School, both Cancer Center Support Grants.

References

1. Tate P. H., Bird A. P. Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genet. Dev., *3:* 226-231, 1993.
2. Jones P. A., Laird P. W. Cancer epigenetics comes of age. Nat. Genet., *21:* 163-167, 1999.
3. Rountree M. R., Bachman K. E., Herman J. G., Baylin S. B. DNA methylation, chromatin inheritance, and cancer. Oncogene, *20:* 3156-3165, 2001.

4. Kovesdi I., Reichel R., Nevins J. R. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. Proc. Natl. Acad. Sci. USA, *84:* 2180-2184, 1987.
5. Iguchi-Ariga S. M., Schaffner W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev., *3:* 612-619, 1989.
6. Watt F., Molloy P. L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev., *2:* 1136-1143, 1988.
7. Graessmann A., Graessmann M. DNA methylation, chromatin structure and regulation of *Herpes simplex* virus *tk* gene expression. Gene (Amst.), *74:* 135-137, 1988.
8. Herman J. G., Umar A., Polyak K., Graff J. R., Ahuja N., Issa J. P., Markowitz S., Willson J. K., Hamilton S. R., Kinzler K. W., Kane M. F., Kolodner R. D., Vogelstein B., Kunkel T. A., Baylin S. B. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc. Natl. Acad. Sci. USA, *95:* 6870-6875, 1998.
9. Nakagawa H., Nuovo G. J., Zervos E. E., Martin E. W., Jr., Salovaara R., Aaltonen L. A., de la Chapelle A. Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. Cancer Res., *61:* 6991-6995, 2001.
10. Kuismanen S. A., Holmberg M. T., Salovaara R., Schweizer P., Aaltonen L. A., de La Chapelle A., Nystrom-Lahti M., Peltomaki P. Epigenetic phenotypes distinguish microsatellite-stable and -unstable colorectal cancers. Proc. Natl. Acad. Sci. USA, *96:* 12661-12666, 1999.
11. Wheeler J. M., Beck N. E., Kim H. C., Tomlinson I. P., Mortensen N. J., Bodmer W. F. Mechanisms of inactivation of mismatch repair genes in human colorectal cancer cell lines: the predominant role of hMLH1. Proc. Natl. Acad. Sci. USA, *96:* 10296-10301, 1999.
12. Syngal S., Fox E., Li C., Dovidio M., Eng C., Kolodner R. D., Garber J. E. Interpretation of genetic test results for hereditary nonpolyposis colon cancer. J. Am. Med. Assoc., *282:* 247-253, 1999.
13. Kolodner R. D., Tytell J. D., Schmeits J., Kane M. F., Das Gupta R., Weger J., Wahlberg S., Fox E. A., Peel D. J., Ziogas A., Garber J. E., Syngal S., Anton-Culver H., Li F. P. Germline MSH6 mutations in colorectal cancer families. Cancer Res., *59:* 5068-5074, 1999.
14. Wahlberg S. S., Schmeits J., Thomas G., Loda M., Garber J. E., Syngal S., Kolodner R. D., Fox E. A. Evaluation of microsatellite instability and immunohistochemistry for the prediction of germline MSH2 and MLH1 mutations in HNPCC families. Cancer Res., *62:* 3485-3492, 2002.
15. Charbonnier F., Olschwang S., Wang Q., Boisson C., Martin C., Buisine M. P., Puisieux A., Frebourg T. MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. Cancer Res., *62:* 848-853, 2002.
16. Yan H., Papadopoulos N., Marra G., Perrera C., Jiricny J., Boland C. R., Lynch H. T., Chadwick R. B., de la Chapelle A., Berg K., Eshleman J. R., Yuan W., Markowitz S., Laken S. J., Lengauer C., Kinzler K. W., Vogelstein B. Conversion of diploidy to haploidy. Nature (Lond.), *403:* 723-724, 2000.
17. Deng G., Chen A., Hong J., Chae H. S., Kim Y. S. Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. Cancer Res., *59:* 2029-2033, 1999.
18. Grady W. M., Rajput A., Lutterbaugh J. D., Markowitz S. D. Detection of aberrantly methylated hMLH1 promoter DNA in the serum of patients with microsatellite unstable colon cancer. Cancer Res., *61:* 900-902, 2001.
19. Kane M. F., Loda M., Gaida G. M., Lipman J., Mishra R., Goldman H., Jessup J. M., Kolodner R. D. 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., *57:* 808-811, 1997.
20. Ito E., Yanagisawa Y., Iwahashi Y., Suzuki Y., Nagasaki H., Akiyama Y., Sugano S., Yuasa Y., Maruyama K. A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene *hMLH1*. Biochem. Biophys. Res. Commun., *256:* 488-494, 1999.
21. Veigl M. L., Kasturi L., Olechnowicz J., Ma A. H., Lutterbaugh J. D., Periyasamy S., Li G. M., Drummond J., Modrich P. L., Sedwick W. D., Markowitz S. D. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc. Natl. Acad. Sci. USA, *95:* 8698-8702, 1998.
22. Deng G., Chen A., Pong E., Kim Y. S. Methylation in hMLH1 promoter interferes with its binding to transcription factor CBF and inhibits gene expression. Oncogene, *20:* 7120-7127, 2001.
23. Cameron E. E., Bachman K. E., Myohanen S., Herman J. G., Baylin S. B. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat. Genet., *21:* 103-107, 1999.
24. Muto S., Horie S., Takahashi S., Tomita K., Kitamura T. Genetic and epigenetic alterations in normal bladder epithelium in patients with metachronous bladder cancer. Cancer Res., *60:* 4021-5402, 2000

Chapter 6

Regulation of the human MSH6 gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms.

Isabella Gazzoli and Richard D. Kolodner

Molecular and Cellular Biology 2003 Nov; 23(22):7992-8007.

Abstract

Defects in human DNA mismatch repair have been reported to underlie a variety of hereditary and sporadic cancer cases. We characterized the structure of the MSH6 promoter region to examine the mechanisms of transcriptional regulation of the *MSH6* gene. The 5′-flanking region of the *MSH6* gene was found to contain seven functional Sp1 transcription factor binding sites that each bind Sp1 and Sp3 and contribute to promoter activity. Transcription did not appear to require a TATA box and resulted in multiple start sites, including two major start sites and at least nine minor start sites. Three common polymorphisms were identified in the promoter region (−557 T→G, −448 G→A, and −159 C→T): the latter two were always associated, and each of these functionally inactivated a different Sp1 site. The polymorphic allele −448 A −159 T was demonstrated to be a common Caucasian polymorphism found in 16% of Caucasians and resulted in a five-Sp1-site promoter that had 50% less promoter activity and was more sensitive to inactivation by DNA methylation than the more common seven Sp1 site promoter allele, which was only partially inactivated by DNA methylation. In cell lines, this five-Sp1-site polymorphism resulted in reduced MSH6 expression at both the mRNA and protein level. An additional 2% of Caucasians contained another polymorphism, −210 C→T, which inactivated a single Sp1 site that also contributes to promoter activity.

Introduction

The human DNA mismatch repair (MMR) system functions to repair mispaired bases in DNA that result from DNA replication errors and thereby prevents the accumulation of mutations due to such replication errors. Biochemical and genetic studies have identified a number of mismatch repair proteins involved in this system, including those encoded by the *MSH2*, *MSH3*, *MSH6*, *MLH1*, *MLH3*, *PMS2*, and *EXO1* genes, as well as the replication proteins PCNA, RFC, RPA, and DNA polymerase delta (for reviews, see references [24] and [33]). Loss of MMR function is associated with both inherited cancer susceptibility and the development of sporadic tumors. Inherited mutations in *MSH2* and *MLH1* are the most prevalent cause of hereditary nonpolyposis colorectal carcinoma (HNPCC) (for a review, see reference [52]), and epigenetic silencing of MLH1 has been found to underlie most MMR defective sporadic cancer cases ([14, 25, 31, 46, 47]). Inherited mutations in *MSH6* have been found in a small proportion (0 to 3%) of HNPCC families and appear to underlie a higher proportion for familial colorectal cancer cases that show later onset and a less pronounced family history than HNPCC ([6, 34, 45, 65, 67, 68]). Mutations in PMS2 have been found in patients with Turcots syndrome but are only rarely found in patients with HNPCC ([11, 39, 64, 66]). Whether or not mutations in *MLH3* or *EXO1* underlie a significant proportion of HNPCC is unclear ([2, 29, 38, 69, 70]).

In the human MMR system, two heterodimeric complexes—MSH2-MSH6 (MutSα) MSH2-MSH3 (MutSβ)—function to recognize mispaired bases in DNA ([1, 19, 21, 49, 61]). MutSα appears to function in the repair of base-base and insertion/deletion mispairs, whereas hMutSβ only appears to function in the repair of insertion/deletion mispairs. In addition, MutSα appears to be relatively more important for the repair of smaller insertion/deletion mispairs, whereas MutSβ is relatively more important for the repair of larger insertion/deletion mispairs ([1, 21, 54]). As a consequence of this partial redundancy, defects in *MSH2*, *MSH3*, and *MSH6* each have different effects: mutations in *MSH2* cause complete loss of MMR, mutations in *MSH6* cause increased accumulation of base substitution mutations, and in some cases single base insertion/deletion mutations and mutations in MSH3 cause increased accumulation of larger insertion/deletion mutations ([42, 57]). Several studies have also shown that alteration in expression levels of the MSH proteins can perturb MMR. Overexpression of human MSH3 inhibits the formation of the MutSα complex, resulting in increased accumulation of mutations ([18, 41, 63]) and in *Saccharomyces cerevisiae* overexpression of *MSH6* results in a mutator phenotype ([8, 15]), although the basis for this is unclear.

Although it is known that alteration in the expression of MSH6 can perturb MMR, only a limited analysis of the structure of the human MSH6 promoter has been performed and no possible regulatory mutations that affect the expression of *MSH6* have been reported. The 5′ untranslated region (UTR) of the human *MSH6* gene has been cloned and sequenced ([1, 60]). The promoter region has a high GC content, and there appear to be multiple start sites for transcription. It is known that human MLH1, PMS1, PMS2, and MSH2 gene promoters share structural characteristic of TATA-less promoters, including multiple start sites and the sporadic presence of CAAT-boxes and GC-boxes, which may indicate a housekeeping function for these genes ([27, 28, 48, 71]). The transcription of a number of mammalian genes is believed to be regulated by GC box target sites for Sp family transcription factors ([58]), in the absence of a TATA-box close to the transcriptional start sites ([3, 4, 72, 73]). The limited information about the MSH6 promoter region suggests that it may be a member of this class of promoters. To better understand the regulation of the *MSH6* gene, we identified seven putative Sp1 binding sites in the MSH6 promoter region and analyzed their role in the regulation of MSH6 expression. We found that the MSH6 promoter is regulated by Sp1 and that all seven Sp1 sites functionally interacted with Sp1. In addition, we identified polymorphic variants of the MSH6 promoter in which different Sp1 sites were inactivated by single-nucleotide polymorphisms (SNPs) resulting in altered promoter activity.

Materials and methods

Cell culture. Chinese hamster ovary (CHO) cells were obtained from the American Type Culture Collection (Manassas, Va.). Human cervical carcinoma (HeLa) cells were provided by Giuseppina Bonizzi (UCSD School of Medicine, La Jolla, Calif.), a human skin fibroblast (BJ) cell line was provided by Jean Wang (UCSD Department of Biology), and human glioblastoma and glioma cell lines were provided by Webster Cavenee (Ludwig Institute, La Jolla, Calif.). Cells were cultured in Dulbecco modified Eagle medium (Gibco-Invitrogen) supplemented with 10% fetal bovine serum (Sigma) and 1% l-glutamine, penicillin, and streptomycin sulfate (Irvine Scientific) at 37°C in a humidified chamber with 5% CO_2. *Drosophila* Schneider SL2 cells were from the American Type Culture Collection and were grown at 26°C under normal atmospheric conditions in Schneider's medium (Gibco-Invitrogen) supplemented with 10% fetal bovine serum.

DNA sequence analysis. For analysis of genomic DNA samples, a 672-bp region of the human *MSH6* gene containing the 5′ UTR and part of the coding region was amplified by PCR with primers 166UF (5′-GTGCCTACTCTATACAAATCTTGAG) and 817LR (5′-

GTGCCTACTCTATACAAATCTTGAG) by using an Advantage GC2 PCR kit (Clontech). PCR was carried out under the following conditions: initial denaturation for 4 min at 94°C; followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, and extension at 68°C for 30 s; followed by a single final cycle extension at 68°C for 4 min. The resulting PCR products were then purified by digestion with *Exo*I and shrimp alkaline phosphatase and sequenced by using a Perkin-Elmer ABI 3700 DNA sequencer (Applied Biosystems) ([34]). Plasmid DNAs were similarly sequenced with plasmid DNA templates purified by using Qiagen Plasmid Midi kits (Qiagen) and primers designed by using the predicted sequence of individual plasmids.

Sequence chromatograms were analyzed to identify base changes, including heterozygous base changes by using Sequencher 3.1 software (Gene Codes Corp.), and sequence homologies were examined by using the BLAST 2.0 and FASTA programs available on the web server of the National Center for Biotechnology Information (Bethesda, Md.) ([51]). The 5′-flanking sequence was analyzed for the presence of putative transcription factor binding sites by using MacVector (Oxford Molecular, Ltd.) and TESS (www.cbil.upenn.edu/cgi-bin/tess) software. The sequence of the 5′ UTR of the *MSH6* gene initially used in these studies was obtained from GenBank and included previous submissions from this and other laboratories (GenBank sequences U73732 and AF334668).

To detect polymorphisms in the 5′ UTR of the *MSH6* gene, normal control DNAs were obtained from Coriell Laboratories (Camden, N.J.) and sequenced as described above. The first 100 samples of the complete human diversity set (catalog no. MPDR450) were sequenced. Subsequently, all 100 DNAs from the 100 Caucasian DNA set (catalog no. HD100CAU) were sequenced.

Genomic DNA and mRNA from the LN299, LN340, and LN443 glioblastoma and LNZ308 glioma cell lines were extracted by using a Puregene DNA isolation kit (Gentra Systems) and the FastTrack mRNA isolation kit (Invitrogen, San Diego, Calif.), respectively. The 5′ UTR and all exons with flanking region of hMHS6 gene were amplified and sequenced by utilizing the primers and conditions described previously ([34]). cDNA was generated by using a First-Strand cDNA synthesis kit (Invitrogen), reverse transcription-PCR was performed with the previously described NS6X1UF and CS6X4LAr.5 primers ([34]), and the resulting PCR fragment including exons 1 and 4 was sequenced. To analyze the linkage between different nucleotide changes, the cDNA amplification product and genomic DNA amplified with primers 166UF and CS6X1L2 ([32]) were cloned into a Topo TA cloning pCR2.1 vector (Invitrogen) and then sequenced by using M13 forward and reverse primers.

Primer extension and RACE (rapid amplification of cDNA ends) assay. Primer extension was performed by using total human adult normal colon RNA (ResGen; Invitrogen Corp.) and a reverse primer 817LR that was complementary to nucleotides 17 to 39 from ATG of *MSH6*. The primer was end labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase, and free nucleotides were removed by using a Microspin G25 column (Amersham-Pharmacia). Primer extension assays were performed in parallel by using Superscript II RNase H$^-$ reverse transcriptase and ThermoScript RNase H$^-$ reverse transcriptase (Invitrogen/Life Technologies). Reactions containing either 5 or 10 μg of total human adult normal colon RNA, 5× buffer (provided by the manufacturer), 10 mM dithiothreitol (DTT), 10 mM deoxynucleoside triphosphates, and reverse primer were denatured at 70°C for 2 min and then incubated at 20 min at 60°C so that the primer could anneal to the RNA template, followed by cooling the reactions on ice for 10 min. Then, GC Melt (final concentration, 1 M), RNasin (final concentration, 0.5 U/μl), and either 48 U of Superscript II RNase H$^-$ reverse transcriptase or 200 U of ThermoScript RNase H$^-$ reverse transcriptase were added so that the final volume was 20 μl; the samples were then incubated at 42°C for 90 min. The reactions were stopped by precipitation with 0.5 M ammonium acetate and 100% ethanol. These samples and molecular weight markers (*Hin*fI digest of X174 DNA; Promega) were electrophoresed in parallel on a 7 M urea-6% polyacrylamide gel run in 0.5× TBE buffer. The radioactive DNA species were then detected by autoradiography, as well as by using a phosphorimager.

5′ RACE was carried out by using a Human Colon Marathon-Ready cDNA amplification kit (Clontech). The primary and nested PCR were performed with the adaptor oligonucleotides AP1 and AP2 (provided in the kit) and a specific reverse primer 817LR or a fluorescent derivative of 817LR. PCR amplification was performed by using an Advantage 2 PCR kit (Clontech) according to the manufacturer's instructions for the Ready Marathon cDNA amplification kit. The fluorescently labeled PCR products were electrophoresed on a Perkin-Elmer ABI 377 DNA sequencer and analyzed by using ABI GeneScan software. In addition, the unlabeled RACE products were cloned into the TA cloning vector pCR2.1 (Invitrogen) by using a TA cloning kit, and 100 independent insert containing clones were isolated and sequenced by using standard M13 and T7 vector primers.

Construction of hMSH6 reporter vectors. To construct a wild-type MSH6 promoter-luciferase reporter plasmid, a 624-bp fragment of MSH6, including 5′-flanking sequences from positions −633 to −9 relative to the ATG was amplified by PCR by using an Advantage GC2 PCR kit (Clontech) as described above except that only 20 cycles of PCR were used. The forward primer 166KUF (5′-GTGCCTACTCTATACAAATCT) contained a *Kpn*I site (5′-GGGGTACC) at its 5′

end and the reverse primer 9BLR (5′-CGGCAAGGCCCAACCGTTC) contained a *Bgl*II site (5′-GAAGATTC) at its 5′ end. The PCR product was digested by with *Kpn*I and *Bgl*II and cloned between the *Kpn*I and *Bgl*II sites of the pGL3 enhancer-luciferase reporter plasmid (Promega), and a fully wild-type clone was identified by DNA sequencing. The full-length of the promoter construct pGL3 −633/−9WT was used to construct all deletion and site-directed mutant derivatives.

A series of 5′ MSH6 promoter deletions (pGL3-490/-9, pGL3-318/-9, pGL3-248/-9, pGL3-248Mut/-9, pGL3-221/-9, pGL3-120/-9) was constructed by amplifying the MSH6 promoter region with different 5′ primers complementary to different regions of MSH6 and containing the 5′ *Kpn*I site indicated above in combination with the reverse primer 9BLR, and these PCR products were inserted between the *Kpn*I and *Bgl*II sites of the pGL3 enhancer-luciferase reporter plasmid. Similarly, 3′ deletion constructs (pGL3-633/-252 and pGL3-633/-193) were generated by using the 166KUF primer and different reverse primers complementary to different regions of MSH6 and containing the above indicated 5′ *Bgl*II site. The first number in the designation of each mutant construct indicates the most N-terminal MSH6 nucleotide present, and the second number indicates the most C-terminal MSH6 nucleotide present.

Plasmids containing individual nucleotide changes were constructed by using different site-directed mutagenesis methods. To construct the pGL3-633/-9+TATA box mutation, a derivative of the forward primer 166TKUF was synthesized to contain a 2-bp difference in the core of the TATA box sequence (TATA→AAAA), and then the plasmid was constructed as described above for the wild-type plasmid. The full-length promoter constructs containing mutations that created some of the promoter SNPs, including pGL3-633/-9 SNP1 (−557G), pGL3-633/-9 SNP2-3 (−448A −159T), and pGL3-633/-9 SNP1-2-3 (−557G −448A −159T), were constructed by amplifying these alleles from genomic DNAs in which these alleles were present and then cloning the resulting PCR product as described for the wild-type plasmid. A single round of overlap extension PCR ([62]) with appropriate mutant primers and pGL3-633/-9WT DNA as a template was used to create other single-mutant plasmids. These single-mutant plasmids included the two promoter SNP-containing plasmids pGL3-633/-9 SNP2 (−448A) and pGL3-633/-9 SNP3 (−159T) and a series of plasmids in which a 2-bp substitution was made to eliminate each individual Sp1 consensus sequence, including pGL3-Sp1-7Mut (CC→AA), pGL3-Sp1-6Mut (GG→TT), pGL3-Sp1-5Mut (CC→AA), pGL3-Sp1-4Mut (CC→AA), pGL3-Sp1-3Mut (GG→TT), and pGL3-Sp1-1/2Mut (CC→AA and GG→TT). (The exact nucleotides changed are indicated in Fig. 1A. The Sp1 sites are numbers from 1 to 7 as indicated in Fig. 1B, with 1 indicating the most N-terminal Sp1 site.) Sequential rounds of overlap extension PCR were then used to generate a series of plasmids containing

different combinations of the above-described mutations, including pGL3-Sp1-7/6Mut, pGL3-Sp1-7/6/5Mut, pGL3-Sp1-7/6/5/4Mut, pGL3-Sp1-7/6/5/4/3Mut, pGL3-Sp1-7/6/5/4/3/2/1Mut, and pGL3-Sp1-7/6/5/4/3/2/1Mut+ TATA box mutation. All plasmids were sequenced to verify that they only contained the desired base changes. Note that the oligonucleotide sequences that were the basis for the 2-bp substitution mutations were also used in the many of the competitor DNAs for the electrophoretic mobility shift assay (EMSA) experiments.

Transient-transfection and luciferase reporter gene expression analyses. HeLa, CHO, and BJ cells (2×10^5) were plated into individual wells of a six-well plate 12 h before transfection, and the medium was changed immediately before transfection. A 1-μg DNA mixture was prepared in 100 μl of OptiMEM I medium (Invitrogen) that contained 0.1 μg of β-galactosidase expression plasmid (pβgal-control; BD Biosciences) as an internal control or 0.4 μg (HeLa cells) or 0.1 μg (CHO/BJ cells) of test plasmid and pUC18 carrier DNA and was then mixed with Fugene in a 1:3 ratio. This mixture was immediately added to one well of cells; after 24 h the cells were washed twice with phosphate-buffered saline, and cell extracts were prepared by resuspending the cells in lysis buffer provided in a luciferase assay kit (Promega). The luciferase activity was then analyzed by mixing 25 μl of cell extract and 100 μl of the luciferase assay substrate (Promega) and measuring the resulting luminescence 10 s in a luminometer (EG&G Berthold Microlumat LB 96P). The luciferase activity was then normalized to the level of β-galactosidase activity present in the same extracts measured by using a β-galactosidase activity kit (Invitrogen) according to the manufacturer's instructions. Each extract was analyzed in duplicate, and at least two independent experiments were performed. The observed luciferase activity reported in each figure is expressed as the percentage of the activity obtained with the most active construct used in each figure. The DNAs and cell lines transfected are as indicated in individual experiments. In some experiments, the plasmids were methylated in vitro with *Sss*I methylase (New England Biolabs) prior to transfection. In these cases, complete methylation at the Sp1 and CpG sites was verified by measuring the extent of protection from digestion with the restriction enzymes *Aci*I (CCGC) and *Hpa*I (CCGG).

For transfection of *Drosophila* SL2 cells, cells were seeded at 2×10^6 cells per well of a six-well plate 24 h before transfection, and transfections were performed and analyzed essentially as described above except that CellFectin reagent (Invitrogen) was used instead of Fugene. The pPac-Vector, pPac-Sp1, and pPac-Sp3 expression plasmids used in the *Drosophila* cell transfection experiments were provided by K. Okumura, (Ludwig Institute for Cancer Research, La Jolla, Calif.). In initial optimization experiments, cotransfection was carried out with different amounts of

pPAC-Sp1 or pPAC-Sp3 (0.1, 0.25, and 0.5 μg), along with either the pGL3-Basic vector or pGL3-633/-9WT, and maximal promoter activity was obtained with 0.25 μg of pPAC-Sp1 and pPAC-Sp3. In subsequent experiments, the DNA transfection mixture contained 0.1 μg of β-galactosidase expression plasmid, 0.4 μg of each experimental plasmid, and 0.25 μg of pPac vector, pPAC-Sp1, or pPAC-Sp3 or a mixture of pPAC-Sp1 and pPAC-Sp3, along with sufficient pUC18 DNA carrier, so that each mixture contained a total of 1 μg of DNA.

Nuclear extracts, EMSAs, and protein analysis. HeLa cells (10^7) were harvested by centrifugation at 6,000 rpm for 5 min, washed twice in ice-cold phosphate-buffered saline, and lysed by resuspending them in 200 μl of ice-cold buffer A (10 mM HEPES [pH 8.0], 10 mM KCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride [PMSF], 0.2% NP-40) and incubating the suspension for 10 min on ice. All solutions contained freshly added protease inhibitor cocktail (1× recommended concentration; Roche) and phosphatase inhibitors (final concentrations of 1 mM orthovanadate and 1 mM sodium fluoride), as well as final concentrations of 1 mM PMSF and 1 mM DTT. After centrifugation at 6,000 rpm at 4°C for 10 min to harvest the nuclei, the nuclei were washed with 200 μl of ice-cold buffer B (10 mM HEPES [pH 8], 10 mM KCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM PMSF) and then resuspended with ice-cold buffer C (20 mM HEPES [pH 8.0], 0.63 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 1 mM PMSF, 25% glycerol) by rotating the suspension for 20 min at 4°C. The resulting lysate was centrifuged at 12,000 rpm for 10 min ([7, 17]), and the protein concentration was determined by using Bio-Rad protein assay reagent.

To construct oligonucleotide duplexes, 500 ng each of sense and antisense oligonucleotides were annealed in buffer N (1 M NaCl, 25 mM Tris-HCl, 1 mM EDTA) by heating the mixtures to 94°C for 10 min and then cooling them to 65°C for 30 min, followed by cooling from 65 to 25°C over a 90-min period, followed by incubation at 25°C for 20 min. The oligonucleotide duplexes were then purified by high-pressure liquid chromatography as previously described, precipitated with ethanol from a solution containing 0.5 M ammonium acetate, and resuspended in 10 μl of TE buffer (10 mM Tris-HCl [pH 8], 1 mM EDTA). A portion (1 μl) of oligonucleotide duplex was then 5′end labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase, and free nucleotides were removed by using Microspin G25 columns (Amersham-Pharmacia). In some cases, the oligonucleotide duplexes were methylated in vitro with *Sss*I methylase, and then the methylation status was verified by digesting the DNAs with the restriction enzyme *Aci*I.

EMSAs was performed by using the gel shift assay system (Promega), with minor modifications. Nuclear extracts (8 μg in final reaction volumes of 20 μl) from HeLa cells were incubated for 20 min at room temperature in the presence or absence of a 50-fold molar excess (relative to the radioactive substrate) of oligonucleotide duplex competitors. Then 20,000 cpm of Sp1 or AP2 consensus oligonucleotide duplexes were added to the reaction and, after 20 min of incubation at room temperature, the samples were analyzed by electrophoresis through a 6% polyacrylamide gel run in 0.5× Tris-borate-EDTA for 2 h at 120 V, followed by detection of the radioactive species by autoradiography or by using a phosphorimager. For the supershift assays, the labeled oligonucleotide duplexes were incubated with the nuclear extract on ice for 20 min as described above; anti-Sp1 (PEP2) or anti-Sp3 (D20) antibody (Santa Cruz Biotechnology) was then added to the reactions on ice for 1 h, and the reaction products were analyzed as described above.

For Western blot analysis, nuclear and cytoplasm extracts from LN229, LN340, and LN443 glioblastoma cell lines were prepared as described above. The nuclear and cytoplasmic extracts from each batch of cells were mixed, and the protein concentration was determined by using a Bio-Rad protein assay kit. A total of 7.5 or 15 μg of protein/sample was electrophoresed through a 4 to 15% sodium dodecyl sulfate-polyacrylamide gel gradient and then transferred to a Hybond-P polyvinylidene difluoride membrane (Amersham Pharmacia Biotech, Inc.). The membrane was probed with monoclonal anti-hMSH6 and anti-hMSH2 (BD Biosciences) antibodies, followed by reaction with horseradish peroxidase-conjugated secondary antibody, and then visualized by using the ECL Plus Light System (Amersham Biosciences). The same membrane was then similarly analyzed by using anti-α tubulin (Sigma) antibody.

Results

Structure of the MSH6 promoter region. To characterize the promoter region of the *MSH6* gene, a fragment of genomic DNA that spans positions −633 to +39 bp relative to the A of the ATG was amplified by PCR and sequenced to confirm the available sequence of this region (Fig. 1A). Analysis of the sequence by using MacVector and TESS software revealed putative binding sites for AP family c-JUN, E2F, E4TF1, and EF.C transcription factors, as well as three potential CAAT boxes and a potential TATA box, identified as a TATAWAW site ([44]) at position −623. No potential Inr elements were found. Also observed were seven potential GC-box/Sp1 transcription factor-binding elements. These included three consensus sequences ([10, 30]) (GGGCGG) at positions −448, −271, and −183 and four inverted complement sequences (CCGCCC) at positions −449,

−235, −208, and −159. Comparison of the sequence of the human MSH6 upstream region with that of other species suggested that the Sp1 site at −159 was conserved in the *MSH6* gene from *Mus musculas*, *Cavia porcellus*, and *Mycobacterium tuberculosis*, whereas the Sp1 site at −271 was conserved in the MSH6 gene from *Bovine herpesvirus typ*.

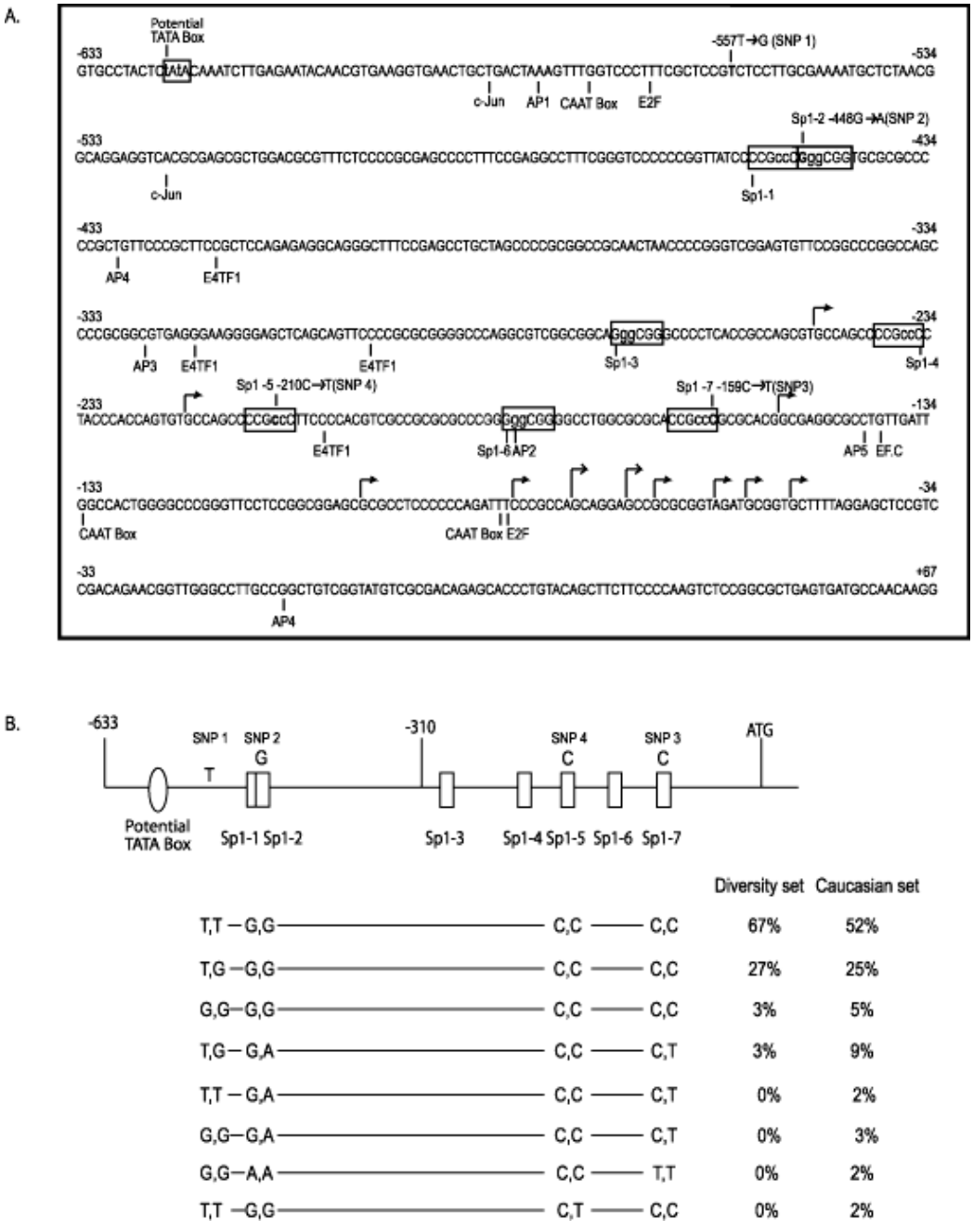

Figure 1. Identification of *cis*-acting elements and SNPs in the MSH6 promoter. **A**. The sequence of nucleotides -633 to -167 relative to the ATG of the *MSH6* gene is shown. The Sp1 sites and the potential TATA box analyzed in the present study are indicated in boxes. Other potential transcription factor binding sites are indicated below the sequence. The major and minor transcription start sites are indicated by the tall and short arrows, respectively. The nucleotides

altered by the 2-bp substitution mutations used in the analysis of the Sp1 sites are indicated in lowercase, and the four SNPs identified are indicated above the sequence. **B**. Schematic representation of the MSH6 promoter region, with the most prevalent nucleotide at each polymorphic site indicated above the promoter diagram. Below the promoter diagram are indicated the eight different genotypes of the MSH6 promoter region identified by analyzing the SNPs present at, from left to right, nucleotides -557 (SNP 1), -448 (SNP 2), -210 (SNP4), and -159 (SNP 3), respectively. Also indicated are the percentages of the samples from the human diversity set of DNAs (first 100 DNAs) and the Caucasian DNAs found to have each genotype.

To identify naturally occurring sequence variations within the promoter region, the region from positions −633 to +39 was amplified and sequenced from the first 100 DNAs of the human DNA diversity set. This analysis revealed three novel SNPs at positions −557 (T→G; SNP 1), −448 (G→A; SNP 2), and −159 (C→T; SNP 3). The frequency of each genotype observed is shown in Fig. 1B. The SNP at −557 was more common than the other two polymorphisms, which were found at the same frequency. Significantly, the −448A and −159T alleles were always associated with each other. The two SNPs at −448A (SNP 2) and −159T (SNP 3) were within the consensus sequences (GGGCGG/CCCGCC) for the potential Sp1 binding sites Sp1-2 and Sp1-7, respectively, and eliminated these two potential Sp1 binding sites.

Subsequent analysis of 100 normal Caucasian DNAs revealed the same three SNPs, except that the frequency of the −448 (G→A; SNP 2) and −159 (C→T; SNP 3) SNPs was much higher in these samples and homozygous −557G −448A −159T and −557T −448A −159T alleles were observed. These results suggest the existence of four haplotypes (−557T −448G −159C; −557G −448G −159C; −557T −448A −159T; and −557G −448A −159T) and further suggest that the −557G −448A −159T and −557T −448A −159T alleles represent common Caucasian polymorphisms that alter the number of Sp1 binding sites present in the MSH6 promoter. In addition, the −210 C→T change (SNP 4) was observed in 2% of the Caucasian samples in association with a −557T −448G −159C haplotype. The −210 C→T change (SNP 4) eliminates the consensus sequence for the putative Sp1-5 site.

Mapping of the MSH6 transcriptional start sites. To map the transcriptional start site(s) of the *MSH6* gene, primer extension analysis of total human normal colon RNA was first performed. This analysis revealed multiple potential 5′ mRNA ends mapping between positions −247 and −51 (data not shown). Because the extremely GC-rich nature of the MSH6 promoter region might cause premature termination of the primer extension products due to secondary structure, 5′-RACE analysis was performed to define the positions of the potential multiple transcription start sites. An anchor primer (AP1) and a specific antisense oligonucleotide were used to amplify Human Colon Marathon-Ready cDNA. Nested PCR was then carried out with internal adaptor primer and fluorescently labeled antisense primer, and the sizes of these PCR products were determined by

using an ABI 377 sequencer. Multiple species of different lengths were observed, and the positions of the potential start sites relative to the position of the antisense primer sequence were determined by subtracting the length of the adapter sequence from the length of each PCR product (Fig. 2).

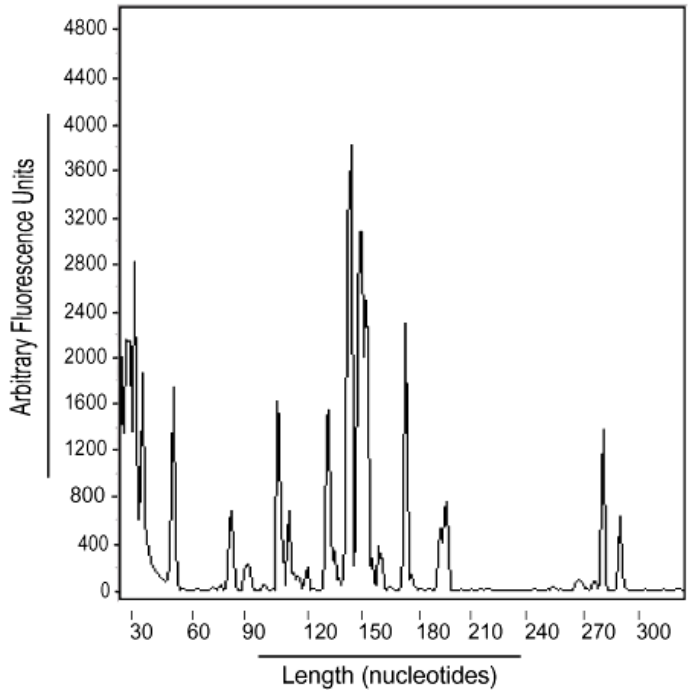

Figure 2. Transcription of the *MSH6* gene initiates from multiple start sites. Genescan analysis of 5-RACE–PCR products was performed as described in Materials and Methods. Below the chromatogram is indicated a molecular size scale in base pairs determined by using appropriate markers (TAMRA).

These positions fell within the range of positions identified by primer extension. Finally, PCR products were generated by using an unlabeled antisense primer, and these were cloned and sequenced to determine the exact position of the start sites. Analysis of 100 clones revealed two major start sites at positions −76 and −70 (found in 49% of the clones), as well as at least nine minor start sites, all of which are indicated in Fig. 1A.

EMSA analysis of the putative Sp1 binding sites. EMSAs were performed to characterize the seven potential Sp1 binding sites identified by computer analysis and to investigate the functional consequences of the SNPs found at −448 (SNP 2) and −159 (SNP 3). Gel shifts were carried out with HeLa cell nuclear extracts as a source of protein and mixtures of one labeled double-stranded oligonucleotide containing the consensus-binding sites for either Sp1 or AP2 and a series of unlabeled double stranded oligonucleotide competitors, including the seven Sp1 sites found between positions −454 and −159, the Sp1 sites containing the −448A and −159T SNPs, and the Sp1 and AP2 consensus sequences (Fig. 3A).

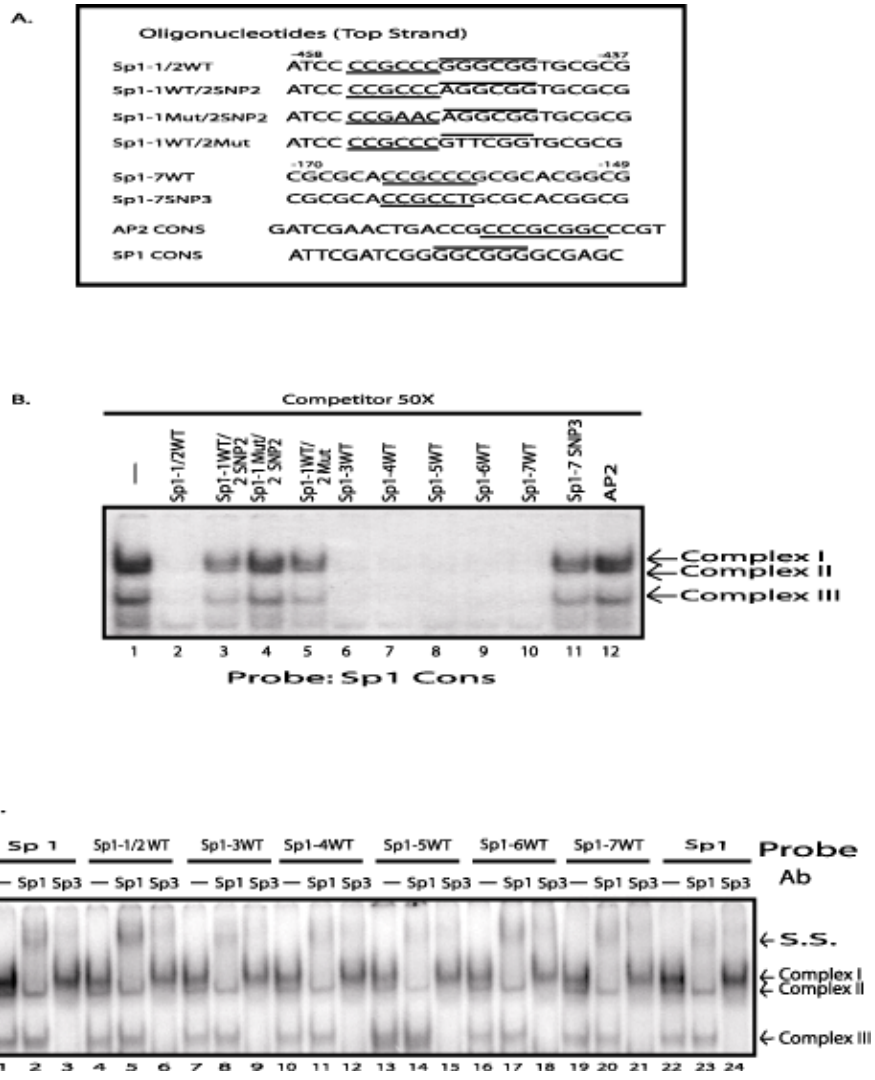

Figure 3. EMSA of the seven Sp1 sites found in the MSH6 promoter region. **A**. Sequences of the competitor and probe DNAs used in gel mobility shift assays. Only the sequence of the top strand of the oligonucleotide duplexes is indicated and, where relevant, the MSH6 sequence coordinates are also indicated. The Sp1 binding consensus sequences (GGGCGG) and inverted complement sequences (CCGCCC) are indicated by the overlining and underlining, respectively. The numbering system for each of the Sp1 sites is as defined in Fig. 1B. Each Sp1 sequence is wild type except that Sp1-1WT/2SNP 2 refers to the sequence containing the _448A SNP 2, Sp1-1Mut/2SNP 2 refers to the sequence containing the 2-bp substitution mutation in the Sp1-1 site and the _448A SNP 2, Sp1-1WT/2Mut refers to the sequence containing the 2-bp substitution mutation in the Sp1-2 site, and Sp1-7SNP 3 refers to the sequence containing the _159T SNP 3. See Fig. 1A for the sequences that are changed by these nucleotide substitutions. **B**. Gel mobility shift assays performed with HeLa cell extract, radioactively labeled Sp1 consensus sequence as a probe, and a 50-fold molar excess of the indicated competitor DNAs. The arrows indicate the three specific protein-DNA complexes formed. **C** Effect of anti-Sp1 and anti-Sp3 antibodies on gel shift assays. Assays were performed with HeLa cell extract, and the radioactive substrate is indicated above each set of three lanes. After complex formation, the antibody indicated above each individual lane was added. The arrows indicate the three specific protein-DNA complexes formed and the position of the supershifted (S.S.) species formed after the addition of antibodies.

When binding to the labeled Sp1 consensus sequence was examined in the absence of competitor, three different protein-DNA complexes were observed (Fig. 3B). A 50-fold excess of competitor DNAs comprising the first two Sp1 sites (Sp1-1/2) or each of the individual Sp1 sites Sp1-3 through Sp1-7 completely eliminated the formation of the three protein DNA complexes (lanes 2 and 6 to 10), whereas the AP2 competitor had no effect on complex formation (lane 12). An Sp1 consensus sequence competitor also eliminated the formation of the three complexes (data not shown). These results suggest that all seven putative Sp1 sites can bind Sp1. The −159T SNP 3 completely eliminated the ability of the Sp1-7 oligonucleotide to act as a competitor of complex formation (lane 11), indicating that this SNP eliminates the Sp1-7 site. The −448A SNP 2 partially eliminated the ability of the Sp1-1/2 oligonucleotide to act as a competitor, and a similar effect was seen with a GG→TT that eliminated the second Sp1 binding site (lanes 3 and 5). However, combining CC→AA mutation eliminating the first Sp1 site with the −448A SNP 2 completely eliminated the ability of the oligonucleotide containing the first two Sp1 sites to act as a competitor of complex formation (lane 4). These latter results suggest that both the Sp1-1 and Sp1-2 sites can bind Sp1 and indicate that the −448A SNP 2 eliminates the Sp1-2 site. As a control, the same unlabeled oligonucleotides were used as competitors in binding reactions containing the labeled AP2 oligonucleotide (data not shown), and competition was only observed with the AP2 competitor and with the Sp1-6 competitor that contains an overlapping AP2 site.

To better characterize the seven Sp1 transcription factor-binding sites, supershift experiments were performed with anti-Sp1 and anti-Sp3 antibodies (Fig. 3C). When the Sp1 consensus sequence and the Sp1-1/2, Sp1-3, Sp1-4, Sp1-5, Sp1-6, or Sp1-7 oligonucleotides were used as labeled substrates in gel shift experiments, the same three protein-DNA complexes were observed (lanes 1, 4, 7, 10, 13, 16, 19, and 22). The addition of an anti-Sp1 antibody to the reactions caused a supershift of only the upper band (complex 1) (lanes 2, 5, 8, 11, 14, 17, 20, and 23), whereas the addition of anti-Sp3 antibody caused a supershift of both of the lower bands (complexes 2 and 3) (lanes 3, 6, 9, 12, 15, 18, 21, and 24). This pattern of supershifting has been observed during the analysis of other promoters whose expression involves Sp1 and Sp3 proteins ([4, 37,36, 58]). These results indicate that complex 1 results from Sp1 binding, that complexes 2 and 3 result from Sp3 binding, and that the Sp1 sites Sp1-1/2, Sp1-3, Sp1-4, Sp1-5, Sp1-6, and Sp1-7 can all bind both Sp1 and Sp3.

Regulation of the MSH6 promoter by Sp1 binding sites. To examine the role of the seven Sp1 binding sites on the activity of the MSH6 promoter, a DNA fragment comprising nucleotides −633 to −9 of MSH6 containing the core promoter was cloned into pGL3 enhancer vector to construct a luciferase reporter vector. When this construct was transiently transfected into either HeLa or CHO

cells, strong luciferase expression was observed compared to transfection of the vector alone (Fig. 4). This MSH6 promoter reporter vector was then used in a series of experiments in which mutations were tested for their effect on expression after transfection into both HeLa and CHO cells; all reported experiments were performed with both cell lines; however, since the same results were obtained with both cell lines, data from only one cell line is presented for each experiment. A 5′ deletion that removed the potential TATA box partially reduced expression; however, a two-base substitution mutation (**T**A**T**A→**A**A**A**A) that eliminated the TATA box had no effect on expression, suggesting the potential TATA box was not important for expression (Fig. 4A). A series of 5′ deletions that progressively eliminated the Sp1-1/2 and Sp1-3 sites resulted in progressive, but not complete, loss of expression. Mutations or deletion of the Sp1-4 site resulted in further reduced expression, and finally deletion of the remaining three Sp1 sites reduced expression to the level seen with the pGL3 vector alone. An internal deletion that eliminated the Sp1-6 and Sp1-7 sites eliminated ca. 80% of the promoter activity, and a deletion that eliminated the Sp1-4, Sp1-5, Sp1-6, and Sp1-7 sites eliminated ca. 90% of the promoter activity. These results suggest that all seven Sp1 sites contribute to the expression of MSH6 but that the four Sp1 sites closest to the transcriptional start sites are most important. The deletion mutations discussed above likely eliminate more than just the seven Sp1 sites. To determine the functional significance of the seven individual Sp1 sites, a series of two-base substitution mutations (G**GG**CGG→G**TT**CGG or CCG**CC**C→CCG**AA**C) eliminating the core Sp1 binding site for each Sp1 site were constructed and tested. The data (Fig. 4B) show that mutation of the individual Sp1-4, Sp1-5, Sp1-6, and Sp1-7 sites significantly reduced MSH6 promoter activity. In this analysis, the Sp1-4 and Sp1-5 mutations appeared to have a somewhat greater effect than the Sp1-6 and Sp1-7 mutations. In contrast, mutations of the Sp1-1/2 and Sp1-3 sites only modestly reduced the promoter activity whereas mutation of the TATA box had little or no effect on promoter activity. A series of successive 2-bp substitution mutations inactivating the seven Sp1 binding sites were then tested to characterize the contribution of all seven Sp1 sites to MSH6 promoter activity (Fig. 4C). Mutation of the Sp1-7 site decreased the expression of luciferase and successive mutation of the Sp1-6 site resulted in a slight increase in the expression of luciferase. This suggests that the presence of the overlapping Ap2 binding site combined with mutation of the Sp1-6 site might result in increased transcription. In contrast, successive mutation of the Sp1-5, Sp1-4, and Sp1-3 resulted in progressive loss of luciferase expression. Mutation of the Sp1-1/2 sites and the TATA box resulted in little if any further loss of luciferase expression; however, this mutant construct still directed luciferase expression at levels which were above that of the vector alone. Clearly, when all of the Sp1 sites were mutated, the TATA box did not drive promoter activity.

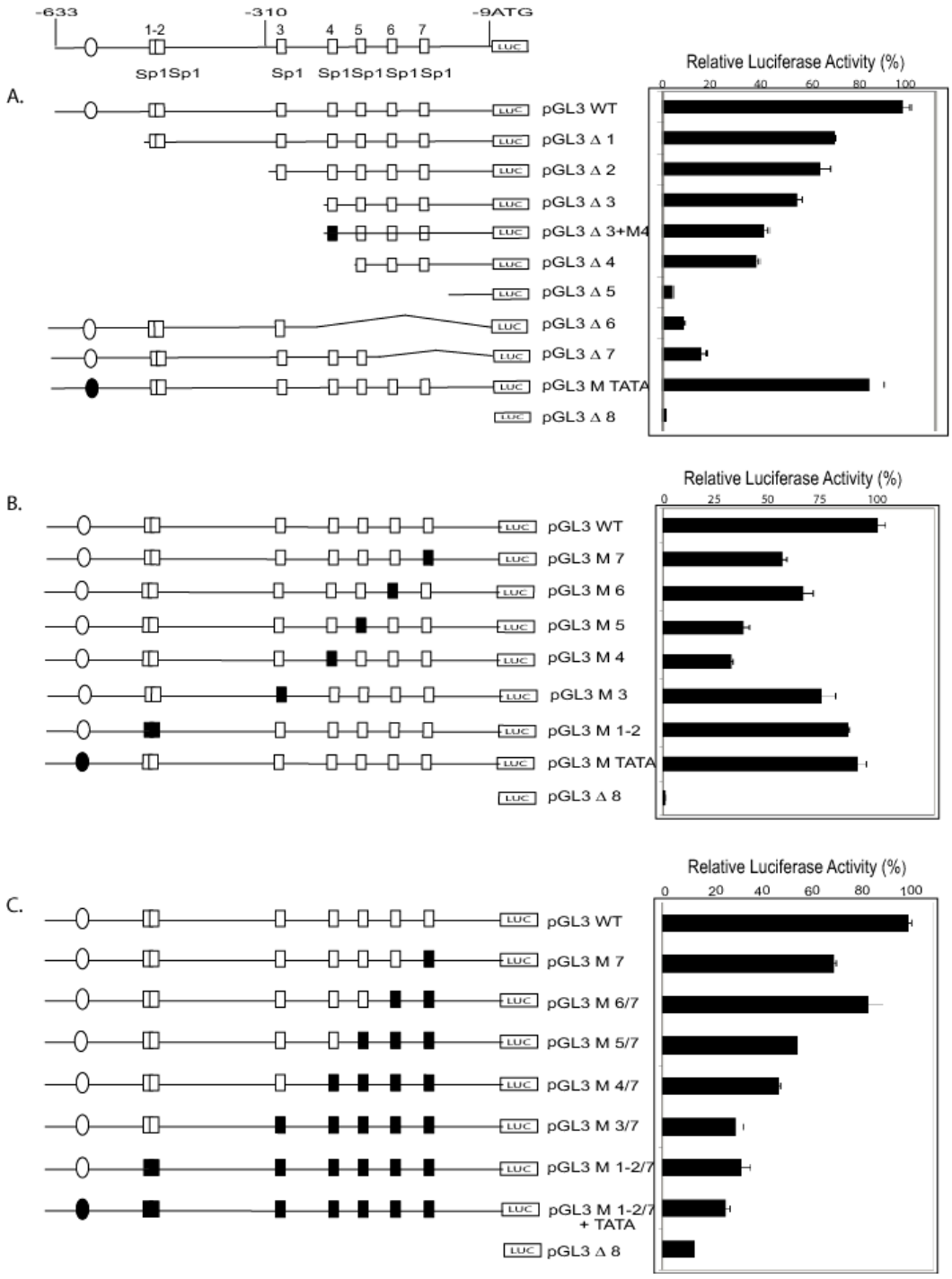

Figure 4. Effects of deletion and two-base substitution mutations on MSH6 promoter activity in transient-transfection assays. The wild-type MSH6 promoter and different mutant derivatives were inserted into the pGL3 enhancer luciferase reporter vector, and the promoter activity of each construct was assayed after transfection into either HeLa or CHO cells. **A**. Analysis of deletion mutant derivatives by transfection into HeLa cells. **B**. Analysis of individual 2-bp substitution mutations after transfection into HeLa cells. **C.** Analysis of combinations of different 2-bp substitution mutations after transfection into CHO cells. The boxes and circles indicate the Sp1 sites and a potential TATA box as indicated in Figure 1, respectively, with the open boxes and circles indicating the wild-type sequences and the solid

boxes and circles indicating the presence of 2-bp substitution mutations: M TATA is a 2-bp substitution in the TATA box, M 1-2 indicates 2-bp substitutions in both the Sp1-1 and Sp1-2 sites, and M 3 through M 7 indicate 2-bp substitutions in the Sp1-3 through Sp1-7 sites, respectively (see Materials and Methods and Fig. 1A for details). The DNA present in the deletion mutants is indicated by the horizontal line and structural features (boxes and circles) present in the diagram: pGL3 WT contains MSH6 nucleotides -633 to -9, pGL3 1 contains MSH6 nucleotides -490 to -9, pGL3 2 contains MSH6 nucleotides -318 to -9, pGL3 3 contains MSH6 nucleotides -248 to -9, pGL3 4 contains MSH6 nucleotides -221 to -9, pGL3 5 contains MSH6 nucleotides -120 to -9, pGL3 6 contains MSH6 nucleotides -633 to -252, pGL3 7 contains MSH6 nucleotides -633 to -193, and pGL3 8 does not contain any MSH6 sequences.

Overall, these results suggest that (i) all seven Sp1 binding sites contribute to MSH6 promoter activity, (ii) that the Sp1-4, Sp1-5, Sp1-6, and Sp1-7 sites are the most important, and (iii) that the promoter does not contain a functional TATA box.

Functional analysis of the three MSH6 promoter SNPs. Two of the three most common SNPs present in the MSH6 promoter region eliminate Sp1 sites and could have an effect on the promoter activity. To investigate this possibility, we generated a series of luciferase reporter constructs containing different combinations of the three SNPs and measured their luciferase expression activity in transient-transfection assays with CHO cells (Fig. 5).

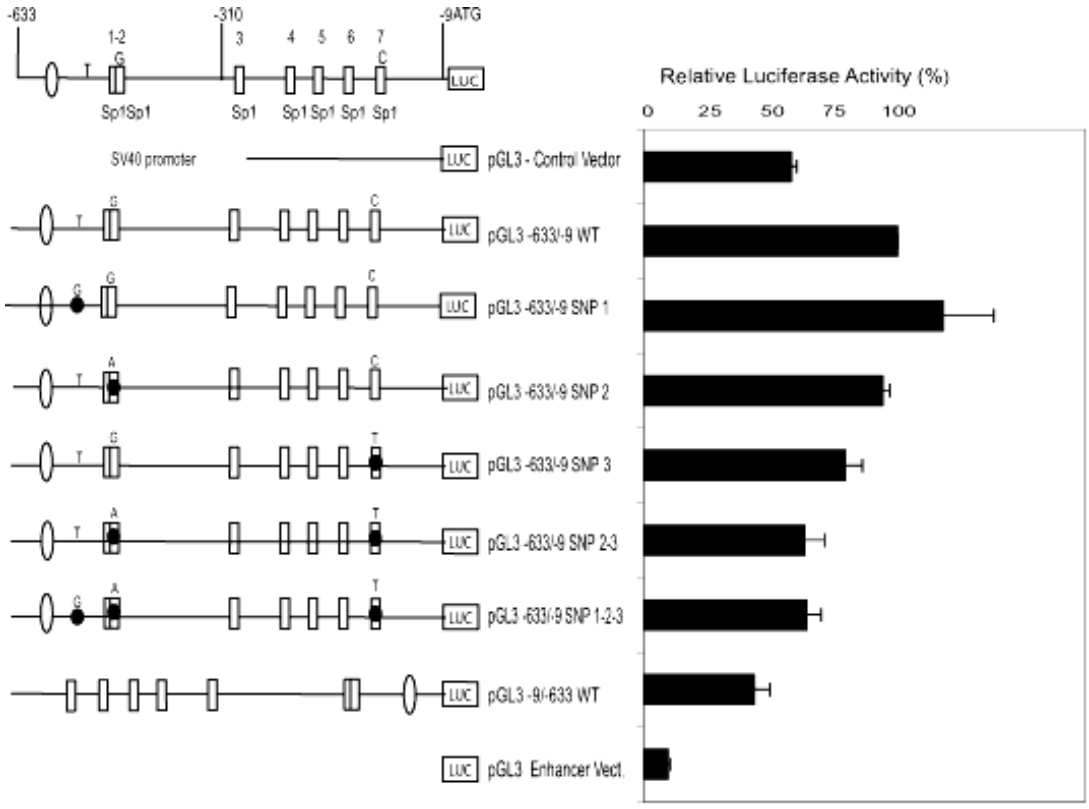

Figure 5. Effect of SNPs on the activity of the MSH6 promoter. The pGL3 wild-type MSH6 promoter reporter vector containing MSH6 nucleotides -633 to -9 and derivatives containing single base substitutions corresponding to the -557G (SNP 1), -448A (SNP 2), and -159T (SNP 3) SNPs or combinations of these substitutions were analyzed for

promoter activity after transfection into CHO cells. The presence of base substitutions of interest is indicated by the nucleotides indicated above each promoter diagram and by the black circle on the relevant structural feature of the promoter diagram when an SNP is present. The pGL3 control vector contains the simian immunodeficiency virus 40 promoter and the vector pGL3-9/-633WT contains the MSH6 promoter region in reverse orientation.

The individual −557G (SNP 1) and −448A (SNP 2) changes did not significantly affect luciferase expression, whereas the single −159T change (SNP 3) reproducibly reduced luciferase expression. These observations are consistent with the above results, indicating that the individual Sp1-7 site is more important for promoter activity than the Sp1-2 site and that the −557 change does not alter a potential transcription factor binding site. The −448A −159T double SNP (SNP 2-3) construct and the −557G −448A −159T triple SNP construct (SNP 1-2-3), the latter of which likely corresponds to the major naturally occurring polymorphic allele, each reproducibly yielded less luciferase expression than the wild-type construct, although expression was higher than that seen for either the promoterless vector or a construct in which the MSH6 promoter was present in reverse orientation. Virtually identical results were obtained when the same constructs were transfected into the human BJ cell line (human skin fibroblasts), whereas the SNPs caused somewhat less reduction of luciferase expression than when the constructs were transfected into HeLa cells.

Transactivation of MSH6 promoter activity by Sp1 and Sp3. To investigate whether Sp1 and Sp3 transfection factors could directly regulate MSH6 promoter activity, *Drosophila* SL2 cells, which are deficient in Sp-related proteins ([59]), were cotransfected with different combinations of Sp1 and Sp3 expression vectors along with different wild-type or mutant MSH6-luciferase reporter constructs (Fig. 6). When the full-length MSH6-luciferase construct or the empty vector (pGL3-Enhancer vector) was cotransfected into SL2 cells, along with increasing amounts of the pPAC-Sp1 and pPAC-Sp3 expression vectors, luciferase expression by the full-length construct was strongly and specifically stimulated by Sp1 expression, whereas Sp3 expression resulted in a lower level of specific expression (Fig. 6A). Consistent with this observation, when different ratios of the Sp1 and Sp3 expression vectors were cotransfected into *Drosophila* SL2 cells, along with the full-length MSH6-luciferase construct or the empty vector, strong MSH6 promoter-dependent luciferase expression was observed with the extent of expression increasing as the ratio of Sp1 to Sp3 increased (Fig. 6B). Finally, the five different SNP containing reporter constructs described above were cotransfected into *Drosophila* SL2 cells, along with either the Sp1 expression construct, the Sp3 expression construct, or a 2:2 ratio of the Sp1 and Sp3 constructs (Fig. 6C). Under the three different Sp transcription factor expression conditions, the −557T (SNP 1) and −448A (SNP 2) single SNPs did not appear to significantly reduce luciferase expression, whereas the −159T SNP (SNP 3), the −448A −159T double SNP (SNP 2-3), and the −557G −448A −159T triple SNP (SNP 1-2-3) combinations all reproducibly reduced luciferase expression.

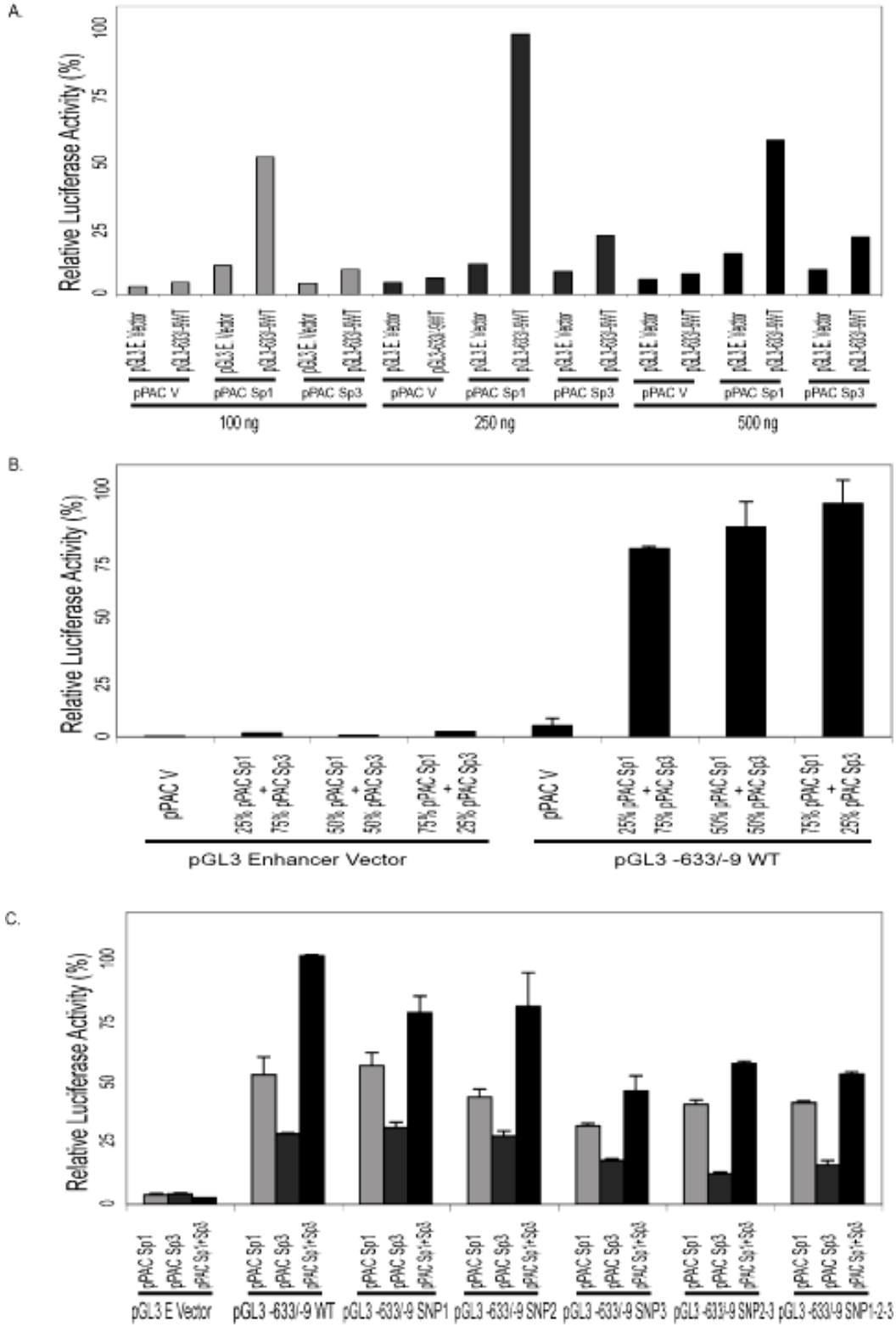

Figure 6. Activation of the MSH6 promoter by Sp1 and Sp3 proteins in SL2 cells. The effect of Sp1 and Sp3 transcription factors on the expression of luciferase driven by the MSH6 promoter was analyzed by cotransfection of the pGL3 vector or the pGL3-633/-9 vector containing the wild-type MSH6 promoter and either Sp1 or Sp3 expression

plasmids into *Drosophila* SL2 cells which lack endogenous Sp1 and Sp3. **A**. *Drosophila* SL2 cells were cotransfected with the indicated amounts of the pPAC-Sp1, pPAC-Sp3, or pPAC (empty) expression vectors and 0.4 μg of either the promoterless reporter construct pGL3-Basic vector or the pGL3-633/-9 vector containing the wild-type MSH6 promoter. **B**. *Drosophila* SL2 cells were cotransfected with 0.4 μg of either the promoterless pGL3-Enhancer vector or the pGL3-633/-9 MSH6 promoter vector and 0.25 μg of either the pPAC empty vector or the indicated mixture of the pPAC-Sp1 and pPAC-Sp3 expression vectors. **C**. *Drosophila* SL2 cells were cotransfected with 0.4 μg of either the promoterless pGL3-Enhancer vector, the pGL3-633/-9 MSH6 promoter vector, or derivatives of pGL3 633/-9 containing the indicated SNPs (see figure 5 for details about the SNP containing plasmids) and 0.25 μg of either the pPAC-Sp1, pPAC-Sp3, or an equimolar mixture of the pPAC-Sp1 and pPAC-Sp3 expression vectors.

These results support the view that the MSH6 promoter is regulated by the Sp transcription factors and that the naturally occurring polymorphic promoter allele containing five Sp1 sites has less promoter activity than the wild-type promoter, which contains seven Sp1 sites.

DNA methylation inhibits Sp1 binding and MSH6 promoter activity. It has previously been reported that DNA methylation can be associated with silencing of MSH6 expression ([5]). Because the Sp1 binding site consensus sequence is a potential substrate for CpG methylation, such methylation-dependent effects could be mediated by methylation of Sp1 sites. To investigate this possibility, the oligonucleotide duplexes containing the Sp1-1/2, Sp1-3, Sp1-4, Sp1-5, Sp1-6, and Sp1-7 sites (Fig. 7A) were methylated in vitro with *Sss*I methylase and tested for their ability to compete with the unmethylated Sp1 consensus site for Sp1/Sp3 binding in EMSA assays (Fig. 7B). Methylation of all six Sp1 site DNAs significantly reduced their ability to compete with the labeled Sp1 consensus oligonucleotide for binding by Sp1 and Sp3 (lanes 3, 5, 7, 9, 11, and 13), although possibly not to the same extent as two-base mutations (G**GG**CGG→G**TT**CGG and CCG**CC**C→CCG**AA**C) that eliminated the Sp1 recognition sequence in each of the MSH6 Sp1 site oligonucleotides (lanes 2, 4, 6, 8, 10, and 12). In comparison, the Sp1 consensus competitor completely eliminated binding. Some previous studies have shown that methylation does not inhibit Sp1 binding ([26]), whereas other studies have found that methylation inhibits Sp1 binding depending on the exact sequence of the Sp1 binding site ([12]); thus, our results suggest that the sequence context of the MSH6 Sp1 sites allows DNA methylation to inhibit Sp1 binding. To further analyze the role of DNA methylation in regulation of the MSH6 promoter, the full-length MSH6 promoter luciferase reporter construct and the two versions containing either the −448A −159T double SNP (SNP 2-3) or the −557G −448A −159T triple SNP (SNP 1-2-3) combinations were treated with *Sss*I methylase. The DNAs were then transfected into HeLa cells (Fig. 7C) or CHO cells (data not shown) to analyze the effect of DNA methylation on luciferase expression. Methylation of the wild-type construct reduced luciferase expression by 90%, whereas methylation of the two SNP-containing constructs reduced luciferase expression to <1% of that seen with the unmethylated wild-type construct.

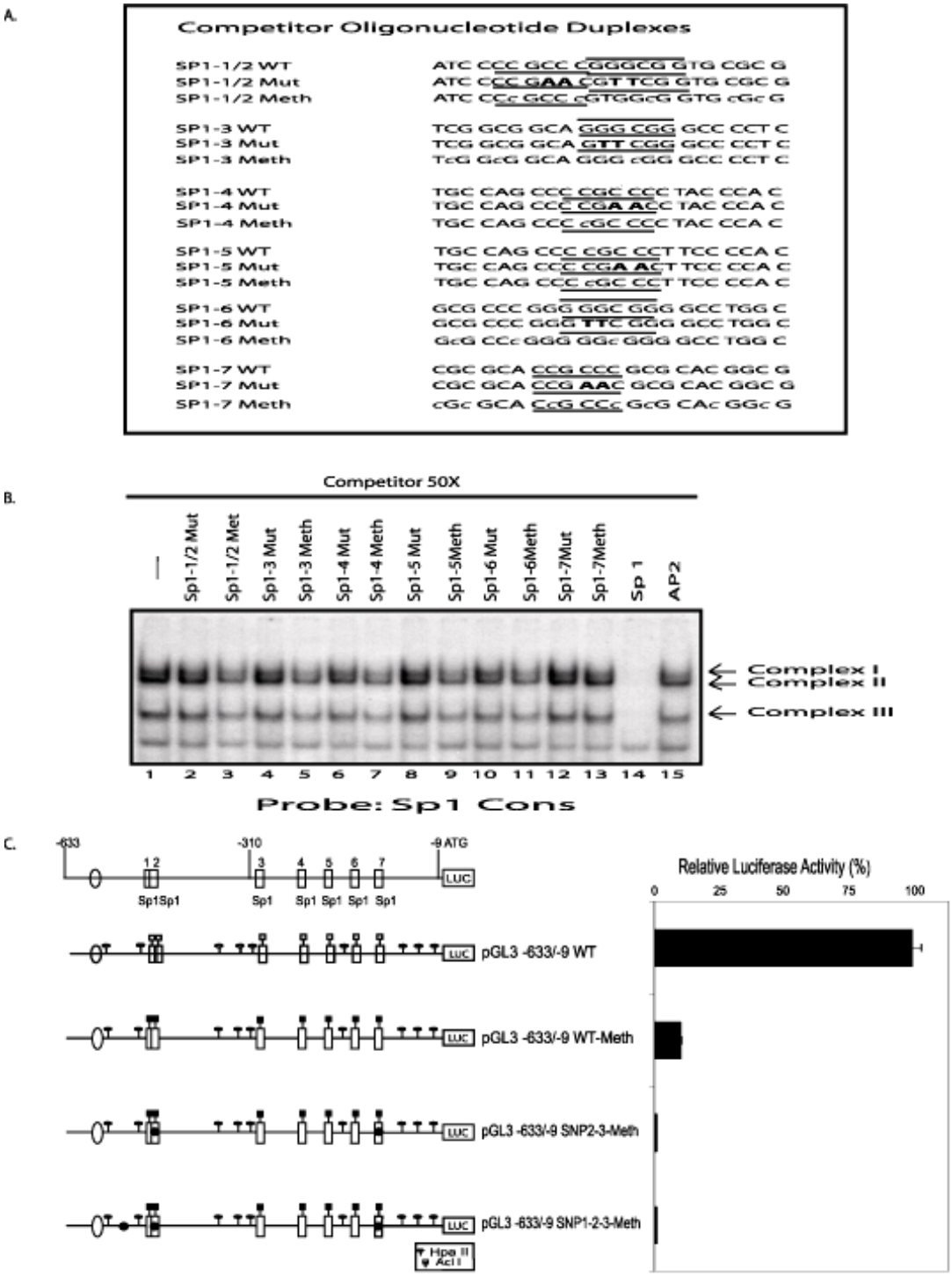

Figure 7. Effect of DNA methylation on the MSH6 promoter Sp1 sites and promoter activity. **A**. Sequences of the competitor DNAs used in gel mobility shift assays. Only the sequence of the top strand of the oligonucleotide duplexes is indicated (see figure 3 for more details). The Sp1 binding consensus sequences (GGGCGG) and inverted complement sequences (CCGCCC) are indicated by the overlining and underlining, respectively. The numbering system for each of the Sp1 sites is as defined in figure 1B. The wild-type sequences are indicated (WT). The mutant

sequence Sp1-1/2Mut refers to the sequence containing the indicated 2-bp substitution mutations in both the Sp1-1 and Sp1-2 sites, whereas otherwise "Mut" refers to the presence of the indicated 2-bp substitution mutations in Sp1 sites Sp1-3 through Sp1-7. See figure 1A for sequences that are changed by these nucleotide substitutions. "Meth" refers to the wild-type sequences that have been methylated in vitro at the Cs designated in lowercase. **B**. Gel mobility shift assays performed with HeLa cell extract, radioactively labeled Sp1 consensus sequence as probe (see figure 3A), and a 50-fold molar excess of the indicated competitor DNAs. The arrows indicate the three specific protein-DNA complexes formed. **C**. The pGL3 wild-type MSH6 promoter reporter vector containing MSH6 nucleotides -633 to -9 and methylated (Meth) versions of this plasmid or derivatives containing single base substitutions corresponding to the -448A (SNP 2) and -159T (SNP 3) SNPs or the -557G (SNP 1), -448A (SNP 2), and -159T (SNP 3) SNPs were analyzed for promoter activity after transfection into HeLa cells. The positions of the *Hpa*II and *Aci*I sites that are substrates for DNA methylation are indicated by the symbols above the line diagram of each promoter, and the presence and position of the SNPs are indicated by the black circles on each line diagram.

These results suggest that the MSH6 promoter could be regulated by methylation, including methylation of its Sp1 binding sites, and that the naturally occurring triple SNP five Sp1 site promoter allele is more sensitive to this type of regulation than the major seven Sp1 site promoter allele.

Analysis of MSH6 expression in tumor cell lines containing MSH6 polymorphisms. To begin to assess the possible in vivo significance of the five Sp1 site promoter, a number of tumor cell lines were screened to identify lines containing different MSH6 polymorphic variants. To do this, the genomic MSH6 locus was sequenced and analyzed for mutations and polymorphisms ([34]). In some cases, fragments of the MSH6 genomic locus and cDNA were amplified by PCR, cloned, and sequenced to determine the linkage between different nucleotide changes present in the promoter, exon 1, and exon 4. Three glioblastoma cell lines were studied. LN229 was found to be homozygous for the seven Sp1 site promoter allele −557T −448G −159C 186C (exon 1) 642C (exon 4) and appears to have at least two copies of MSH6, as evidenced by the presence of heterozygous nucleotides at the site of common silent intronic polymorphisms ([34]). LN340 was found to be heterozygous for the five and seven Sp1 site alleles and for two silent coding sequence polymorphisms; one allele was −557G −448G −159C 186A (exon 1) 642C (exon 4), and the other allele was −557G −448A −159T 186C (exon 1) 642T (exon 4). LN443 only contained the five Sp1 site allele −557G −448A −159T 186C (exon 1) 642T (exon 4). This cell line was probably hemizygous because no heterozygous nucleotides were observed at the site of common silent intronic polymorphisms. None of the cell lines contained a loss-of-function MSH6 mutation or amino-acid-changing variant.

The cell lines LN229, LN340, and LN443 were first analyzed by Western blotting with antibodies to detect MSH6, MSH2, and tubulin as a loading control (Fig. 8A). LN340, the cell line that is heterozygous for the five and seven Sp1 site promoters, had reduced MSH6 levels compared to that of LN229, the cell line that was homozygous for the seven Sp1 site promoter. LN443, the cell line

that was homozygous for the five Sp1 site promoter, had further reduced very low, but still detectable, levels of MSH6. All three cell lines expressed similar levels of MSH2 and tubulin. Treatment of LN340 and LN443 cells with 5-aza-2′-deoxycytidine did not increase the MSH6 protein levels, suggesting that the reduced MSH6 expression was not due to DNA methylation associated silencing in these cell lines. These results are consistent with the idea that the five Sp1 site promoter is less active than the seven Sp1 site promoter in vivo.

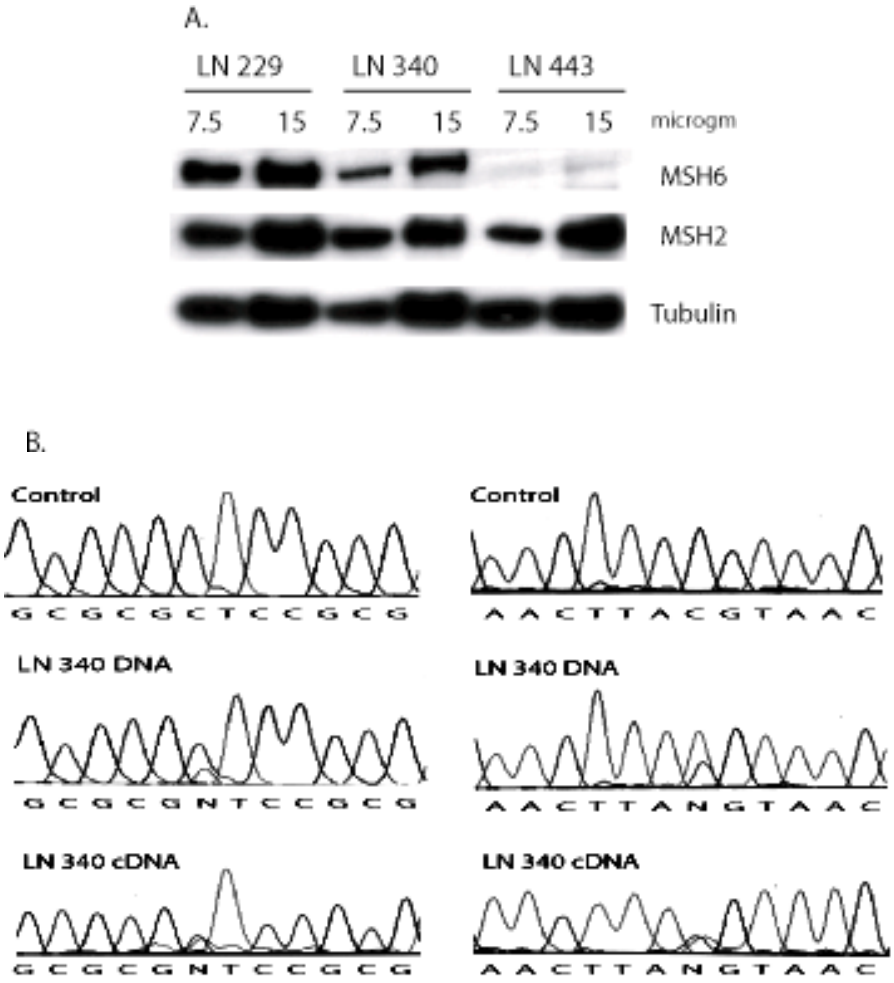

Figure 8. Analysis of MSH6 expression in glioblastoma cell lines. **A**. Western blot analysis of MSH6, MSH2, and tubulin expression. Cell extracts were prepared from the LN229, LN340, and LN443 and analyzed by Western blotting as described in Materials and Methods. Portions (7.5 and 15 μg) of each extract were analyzed as indicated above the individual lanes. The MSH6, MSH2, and tubulin bands are indicated on the right side of the gel. **B**. Sequence analysis of MSH6 mRNA expression from the five and seven Sp1 site promoters. Genomic DNA and cDNA from LN340 cells were prepared and sequenced as described in Materials and Methods. LN340 cells are heterozygous for two different MSH6 alleles: one allele is the seven Sp1 site promoter allele -557G -448G -159C 186A (exon 1) 642C (exon 4), and the other allele is the five Sp1 site promoter allele -557G -448A -159T 186C (exon 1) 642T (exon 4). The sequences on the left cover the region around nucleotide 186 from a control (186C) DNA and LN340 genomic and cDNA as indicated. The sequences on the right cover the region around nucleotide 642 from a control (642C) DNA and LN340

genomic and cDNA as indicated. Reduction (high peak) of the relative levels of the 186C and 642T nucleotides linked to the five Sp1 site promoter compared to the levels of the 186A and 642C nucleotides linked to the seven Sp1 site promoter is seen in the cDNA relative to the genomic DNA.

To further analyze the activity of the five Sp1 site promoter, mRNA was isolated from LN340 cells and used as a template for synthesis of cDNA. Then the region around nucleotides 186 and 642 was amplified by PCR from cDNA and genomic DNA and sequenced to determine the relative amounts of the two polymorphic nucleotides present at each site in each sample (Fig. 8B). Compared to the genomic DNA control, analysis of the relative peak heights in the cDNA indicated that the 186C and 642T variants that are linked to the five Sp1 site promoter were reduced by ca. 60% relative to the 186A and 642C variants linked to the seven Sp1 site promoter. In a control experiment with the LNZ308 cell line that is homozygous for the seven Sp1 site allele and heterozygous for the 186C/A (exon 1) polymorphism, there was no difference between the relative amounts of the 186C and 186A alleles when genomic DNA and cDNA were similarly analyzed (data not shown). Similar to the Western blotting analysis, these results support the view that the five Sp1 site promoter is less active than the seven Sp1 site promoter, a finding consistent with the conclusion from the transfection analysis that the −448A and −159T polymorphisms each inactivate an Sp1 site that is important for the full activity of the promoter. We are currently performing a detailed molecular analysis of the consequences of reduced MSH6 expression in these cell lines.

Discussion

In the present study, we investigated the structure of the human MSH6 core promoter located within a DNA fragment containing nucleotides −633 to +1 relative to the MSH6 translational start site. Computer analysis indicated the presence of seven GC boxes potentially capable of binding Sp1 and Sp3 transcription factors located between positions −454 and −159, as well as other potential transcription factor binding sequences, including AP family, c-JUN, E2F, E4TF1, EF.C, and CAAT box sequences. All seven of the GC boxes partially contributed to promoter activity and were able to bind Sp1 and Sp3. MSH6 transcription appeared to initiate from multiple start sites located between nucleotides −250 to −50 and, consistent with this, deletion analysis showed that a potential TATA box at −623 was not required for transcription. These features are similar to those reported for other human MMR and MMR-related genes such as MSH2 and MLH1 ([27, 28]), where multiple start sites and the absence of a functional TATA box have been reported, and for PMS1 and PMS2, where the promoter regions have been reported to contain CAAT boxes, CpG islands, and Sp1 binding sites ([48, 71]). TATA-less promoters are often activated by Sp family proteins.

Similar to the results reported here for MSH6, Sp1-dependent promoters contain multiple GC boxes located within several hundred base pairs upstream of multiple transcription start sites, and these GC boxes are recognized by the Sp1 protein ([20]), which actives mRNA synthesis by RNA polymerase II ([13]). In this respect, MMR genes, including *MSH6*, appear to be housekeeping genes. As part of the present study, we identified two different polymorphic alleles of MSH6 that had either five or six functional Sp1 sites and provided evidence that the five Sp1 site promoter is less active in vivo than the seven Sp1 site promoter, further supporting the view that MSH6 is regulated by Sp1. This suggests that the human population contains individuals with different levels of MSH6 promoter activity.

To characterize the transcriptional regulation of the *MSH6* gene, we performed a functional analysis of the 5′-flanking region of the gene. The mapping of the 5′ ends of the MSH6 mRNA by 5′-RACE and primer extension assays revealed multiple start sites, including two major start sites that mapped between −76 and −70 bp from ATG. This finding confirmed the previous mapping of MSH6 start sites by using an RNase protection assay ([60]), although we observed a greater number of potential start sites than previously reported and mapped them to the nucleotide level. Seven potential Sp1 binding sites were observed in the proximity of the multiple transcription start sites, which suggests that the MSH6 promoter might be regulated by Sp1 and Sp3 transcription factors. To investigate this possibility, a series of deletion and two-base substitution mutations altering these Sp1 sites were tested for their effect on expression of a reporter construct in transient-transfection assays. This analysis demonstrated that all seven Sp1 sites contributed to maximal expression and also suggested that the first four Sp1 sites upstream of the ATG were more important for expression than the three most upstream Sp1 sites. This type of regulation by multiple Sp1 sites has been observed for numerous other genes ([23, 43, 50]) for which synergistic activation due to binding of multiple Sp1 proteins is important for transcription and regulation of gene expression. Consistent with the view that MSH6 is an Sp1-regulated promoter, all seven Sp1 sites were found to bind both Sp1 and Sp3 and expression of an MSH6 core promoter construct in *Drosophila* SL2 cells that lack Sp1 and Sp3 proteins was absolutely dependent on cotransfection with either Sp1 or Sp3 expression constructs. In addition, the one potential TATA box identified in the MSH6 promoter region was not required for expression of MSH6, a finding consistent with that observed with other Sp1-regulated genes. Overall, the MSH6 promoter region resembles that of other MMR genes and related genes, such as *MSH2*, *MLH1*, *PMS1*, and *PMS2*, and is consistent with the idea that these genes are typical housekeeping genes that are transcribed by RNA polymerase II.

The presence of multiple Sp1 sites in the MSH6 promoter has a number of implications for the regulation of MSH6 expression. It is known that Sp1 and Sp3 can act as transcriptional activators or repressors, depending on the specific promoter involved and also on the form of Sp3 (short or long) present in the cell ([3, 16, 23, 53]). In addition, the hypothetical model of Kwon et al. ([35]) has suggested that the number of sites occupied by Sp1 or Sp3, which compete for the same binding site, are determined by levels and the relative proportion of Sp1 and Sp3 in the cell and affect both transcription initiation and promoter activation. The cotransfection analysis with the *Drosophila* SL2 cells presented here indicates that MSH6 is more strongly activated by Sp1 than by Sp3. This suggests that MSH6 expression could be affected by the relative levels of these two transcription factors present in different cell types.

Analysis of the promoter of other MMR genes and of related genes such as MLH1, MSH2, and PMS1 has identified polymorphisms in the 5′ upstream of these genes that did not appear to affect the transcription of the genes ([27, 28,71]). Recently, there has been a report of germ line mutations in the MHS2 promoter that may have pathogenic effects in some suspected HNPCC and sporadic colorectal cancer patients ([56]). To investigate the presence of polymorphisms in the MSH6 promoter region, nucleotides −633 to +39 were sequenced from 100 samples of the human diversity set of normal control DNAs and from 100 normal Caucasian control DNAs. Compared to the wild-type sequence (−557T −448G −159C), the most common polymorphic allele was −557G −448G −159C, which was heterozygous or homozygous in 30% of the diversity and Caucasian sample sets. The −557G (SNP 1) change does not appear to alter any sequence required for transcription factor binding and, consistent with this, the −557G −448G −159C allele did not cause altered promoter activity in transfection experiments. Interestingly, 16% of the Caucasian DNAs were heterozygous or homozygous for the −557G −448A −159T or −557T −448A −159T alleles, which were present in the diversity set at a lower frequency, a finding consistent with the proportion of Caucasian DNAs in the diversity set. The −448A (SNP 2) and −159T (SNP 3) changes eliminated the Sp1-2 and Sp1-7 consensus sites and, a finding consistent with this, the mutant sites were no longer bound by Sp1 or Sp3. In transfection experiments, the −557G −448A −159T allele showed a 50% reduction in promoter activity, and this allele was significantly more sensitive to inactivation by DNA methylation than the wild-type allele. Analysis of MSH6 expression in cell lines containing the −557G −448A −159T allele showed that the five Sp1 site promoter was less active than the seven Sp1 site promoter, resulting in reduced MSH6 expression at both the mRNA and protein level; this result provides further evidence that MSH6 is a Sp1 regulated gene in vivo. We also observed that 2% of the Caucasian DNAs were heterozygous for the −210T change (SNP 4); these data and two additional examples (not shown) are consistent with the polymorphic allele

being −557T −448G −210T −159C. The −210T change (SNP 4) is predicted to inactivate the Sp1-5 site, which is particularly important for promoter activity, although this polymorphic variant was not analyzed directly for its effect on promoter activity. These results suggest that Caucasians contain two relatively common polymorphic variants that result in reduced MSH6 promoter activity. However, at present, we have no evidence that either variant is associated with increased cancer susceptibility. Finally, we observed five other SNPs at low frequencies in the diversity set but not the Caucasian set, none of which appear to alter sequences predicted to be important for promoter activity.

The MSH6 promoter region has a high GC content, a finding consistent with the presence of seven Sp1 sites in the promoter region. This raises the possibility that the MSH6 promoter region could be a target of DNA methylation at CpG sites ([12, 22]). No studies have yet examined MSH6 expression and methylation in human tumors; however, a human tumor cell line has been reported in which MSH6 expression was absent but could be reactivated by treatment with 5-azacytidine, indicating that MSH6 expression can be eliminated by DNA methylation associated gene silencing ([5]). Consistent with this, we have observed that methylation of the MSH6 promoter region at CpG sites by *Sss*I methylase significantly reduces promoter activity. In addition, consistent with some published reports ([12]), methylation of oligonucleotides containing the different MSH6 Sp1 sites reduced the ability of each site to bind Sp1 and Sp3, although not as much as a mutation eliminating each Sp1 site. This suggests that methylation of individual Sp1 sites may result in reduced promoter activity by reducing Sp1 binding. In addition, it is thought ([12]) that binding of Sp1 transcription factors to their GC-rich binding sites helps maintain CpG islands in an unmethylated state. As a consequence, mutation or deletion of Sp1 sites could reduce binding of Sp1 transcription factors, allowing de novo methylation of CpG islands ([9, 32, 40]). This has been reported for the Rb gene, where mutations resulted in promoter silencing ([55]). This finding suggests that, because the two polymorphisms (−448A, SNP 2; −159T, SNP 3) decrease the amount of Sp1 bound to the MSH6 promoter and hence increase the sensitivity of the MSH6 promoter to silencing by DNA methylation in vitro, this polymorphic allele could be a preferential target of de novo methylation, resulting in transcriptional silencing of the gene compared to the more prevalent seven Sp1 site promoter. It is also possible that reduced transcription due to methylation of Sp1 sites could then facilitate the binding of other factors to methylated DNA, resulting in more complete silencing of MSH6. However, to our knowledge, there have been no reports of epigenetic silencing of MSH6 in tumors, although there have been reports of loss of expression of MSH6 in tumors.

Acknowledgments

We are grateful to K. Arden, G. Bonizzi, W. Cavenee, A. C. Li, and J. Wang for providing cell lines and for technical support. We thank J. Weger and J. Mueller for performing DNA sequencing and A. Rojas and T. Sternsdorf for help with sequence analysis software. Finally, we thank C. Glass, B. Guerra, H. Hoffman, A. R. Houweling, and K. Okumura for helpful discussions and comments on the manuscript. This work was supported by NIH grant ES11040.

References

1. Acharya, S., T. Wilson, S. Gradia, M. F. Kane, S. Guerrette, G. T. Marsischky, R. Kolodner, and R. Fishel. 1996. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proc. Natl. Acad. Sci. USA 93:13629-13634.
2. Akiyama, Y., H. Nagasaki, T. Nakajima, H. Sakai, T. Nomizu, and Y. Yuasa. 2001. Infrequent frameshift mutations in the simple repeat sequences of hMLH3 in hereditary nonpolyposis colorectal cancers. Jpn. J. Clin. Oncol. 31:61-64.
3. Andrew, S. D., P. J. Delhanty, L. M. Mulligan, and B. G. Robinson. 2000. Sp1 and Sp3 transactivate the RET proto-oncogene promoter. Gene 256:283-291.
4. Aslam, F., L. Palumbo, L. H. Augenlicht, and A. Velcich. 2001. The Sp family of transcription factors in the regulation of the human and mouse MUC2 gene promoters. Cancer Res. 61:570-576.
5. Bearzatto, A., M. Szadkowski, P. Macpherson, J. Jiricny, and P. Karran. 2000. Epigenetic regulation of the MGMT and hMSH6 DNA repair genes in cells resistant to methylating agents. Cancer Res. 60:3262-3270.
6. Berends, M. J., Y. Wu, R. H. Sijmons, R. G. Mensink, T. van der Sluis, J. M. Hordijk-Hos, E. G. de Vries, H. Hollema, A. Karrenbeld, C. H. Buys, A. G. van der Zee, R. M. Hofstra, and J. H. Kleibeuker. 2002. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. Am. J. Hum. Genet. 70:26-37.
7. Bonizzi, G., J. Piette, S. Schoonbroodt, M. P. Merville, and V. Bours. 1999. Role of the protein kinase C lambda/iota isoform in nuclear factor-kappaB activation by interleukin-1beta or tumor necrosis factor-alpha: cell type specificities. Biochem. Pharmacol. 57:713-720.
8. Bowers, J., T. Sokolsky, T. Quach, and E. Alani. 1999. A mutation in the MSH6 subunit of the *Saccharomyces cerevisiae* MSH2-MSH6 complex disrupts mismatch recognition. J. Biol. Chem. 274:16115-16125.
9. Brandeis, M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, A. Nemes, V. Temper, A. Razin, and H. Cedar. 1994. Sp1 elements protect a CpG island from de novo methylation. Nature 371:435-438.
10. Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science 234:47-52.
11. Chadwick, R. B., J. E. Meek, T. W. Prior, P. Peltomaki, and A. de La Chapelle. 2000. Polymorphisms in a pseudogene highly homologous to PMS2. Hum. Mutat. 16:530.
12. Clark, S. J., J. Harrison, and P. L. Molloy. 1997. Sp1 binding is inhibited by (m)Cp(m)CpG methylation. Gene 195:67-71.
13. Courey, A. J., and R. Tjian. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell 55:887-898.
14. Cunningham, J. M., E. R. Christensen, D. J. Tester, C. Y. Kim, P. C. Roche, L. J. Burgart, and S. N. Thibodeau. 1998. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res. 58:3455-3460.
15. Das Gupta, R., and R. D. Kolodner. 2000. Novel dominant mutations in *Saccharomyces cerevisiae* MSH6. Nat. Genet. 24:53-56.

16. Dennig, J., M. Beato, and G. Suske. 1996. An inhibitor domain in Sp3 regulates its glutamine-rich activation domains. EMBO J. 15:5659-5667.
17. Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11:1475-1489.
18. Drummond, J. T., J. Genschel, E. Wolf, and P. Modrich. 1997. DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSα/hMutSβ ratio and reduces the efficiency of base-base mismatch repair. Proc. Natl. Acad. Sci. USA 94:10144-10149.
19. Drummond, J. T., G. M. Li, M. J. Longley, and P. Modrich. 1995. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science 268:1909-1912.
20. Dynan, W. S., and R. Tjian. 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316:774-778.
21. Genschel, J., S. J. Littman, J. T. Drummond, and P. Modrich. 1998. Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. J. Biol. Chem. 273:19895-19901.
22. Graff, J. R., J. G. Herman, S. Myohanen, S. B. Baylin, and P. M. Vertino. 1997. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. J. Biol. Chem. 272:22322-22329.
23. Hagen, G., S. Muller, M. Beato, and G. Suske. 1994. Sp1-mediated transcriptional activation is repressed by Sp3. EMBO J. 13:3843-3851.
24. Harfe, B. D., and S. Jinks-Robertson. 2000. DNA mismatch repair and genetic instability. Annu. Rev. Genet. 34:359-399.
25. Herman, J. G., A. Umar, K. Polyak, J. R. Graff, N. Ahuja, J. P. Issa, S. Markowitz, J. K. Willson, S. R. Hamilton, K. W. Kinzler, M. F. Kane, R. D. Kolodner, B. Vogelstein, T. A. Kunkel, and S. B. Baylin. 1998. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc. Natl. Acad. Sci. USA 95:6870-6875.
26. Holler, M., G. Westin, J. Jiricny, and W. Schaffner. 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. Genes Dev. 2:1127-1135.
27. Ito, E., Y. Yanagisawa, Y. Iwahashi, Y. Suzuki, H. Nagasaki, Y. Akiyama, S. Sugano, Y. Yuasa, and K. Maruyama. 1999. A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene hMLH1. Biochem. Biophys. Res. Commun. 256:488-494.
28. Iwahashi, Y., E. Ito, Y. Yanagisawa, Y. Akiyama, Y. Yuasa, T. Onodera, and K. Maruyama. 1998. Promoter analysis of the human mismatch repair gene hMSH2. Gene 213:141-147.
29. Jagmohan-Changur, S., T. Poikonen, S. Vilkki, V. Launonen, F. Wikman, T. F. Orntoft, P. Moller, H. Vasen, C. Tops, R. D. Kolodner, J. P. Mecklin, H. Jarvinen, S. Bevan, R. S. Houlston, L. A. Aaltonen, R. Fodde, J. Wijnen, and A. Karhu. 2003. EXO1 variants occur commonly in normal population: evidence against a role in hereditary nonpolyposis colorectal cancer. Cancer Res. 63:154-158.
30. Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51:1079-1090.
31. Kane, M. F., M. Loda, G. M. Gaida, J. Lipman, R. Mishra, H. Goldman, J. M. Jessup, and R. Kolodner. 1997. 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. 57:808-811.
32. Knight, S. W., T. J. Vulliamy, B. Morgan, K. Devriendt, P. J. Mason, and I. Dokal. 2001. Identification of novel DKC1 mutations in patients with dyskeratosis congenita: implications for pathophysiology and diagnosis. Hum. Genet. 108:299-303.
33. Kolodner, R. D., and G. T. Marsischky. 1999. Eukaryotic DNA mismatch repair. Curr. Opin. Genet. Dev. 9:89-96.
34. Kolodner, R. D., J. D. Tytell, J. L. Schmeits, M. F. Kane, R. D. Gupta, J. Weger, S. Wahlberg, E. A. Fox, D. Peel, A. Ziogas, J. E. Garber, S. Syngal, H. Anton-Culver, and F. P. Li. 1999. Germ-line *msh6* mutations in colorectal cancer families. Cancer Res. 59:5068-5074.
35. Kwon, H. S., M. S. Kim, H. J. Edenberg, and M. W. Hur. 1999. Sp3 and Sp4 can repress transcription by competing with Sp1 for the core *cis*-elements on the human ADH5/FDH minimal promoter. J. Biol. Chem. 274:20-28.
36. Lania, L., B. Majello, and P. De Luca. 1997. Transcriptional regulation by the Sp family proteins. Int. J. Biochem. Cell Biol. 29:1313-1323.
37. LeVan, T. D., J. W. Bloom, T. J. Bailey, C. L. Karp, M. Halonen, F. D. Martinez, and D. Vercelli. 2001. A common single nucleotide polymorphism in the CD14 promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. J. Immunol. 167:5838-5844.
38. Lipkin, S. M., V. Wang, D. L. Stoler, G. R. Anderson, I. Kirsch, D. Hadley, H. T. Lynch, and F. S. Collins. 2001. Germline and somatic mutation analyses in the DNA mismatch repair gene MLH3: evidence for somatic mutation in colorectal cancers. Hum. Mutat. 17:389-396.

39. Liu, T., H. Yan, S. Kuismanen, A. Percesepe, M. L. Bisgaard, M. Pedroni, P. Benatti, K. W. Kinzler, B. Vogelstein, M. Ponz de Leon, P. Peltomaki, and A. Lindblom. 2001. The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. Cancer Res. 61:7798-7802.
40. Macleod, D., J. Charlton, J. Mullins, and A. P. Bird. 1994. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev. 8:2282-2292.
41. Marra, G., I. Iaccarino, T. Lettieri, G. Roscilli, P. Delmastro, and J. Jiricny. 1998. Mismatch repair deficiency associated with overexpression of the MSH3 gene. Proc. Natl. Acad. Sci. USA 95:8568-8573.
42. Marsischky, G. T., N. Filosi, M. F. Kane, and R. Kolodner. 1996. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev. 10:407-420.
43. Mastrangelo, I. A., A. J. Courey, J. S. Wall, S. P. Jackson, and P. V. Hough. 1991. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. Proc. Natl. Acad. Sci. USA 88:5670-5674.
44. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science 217:316-324.
45. Miyaki, M., M. Konishi, K. Tanaka, R. Kikuchi-Yanoshita, M. Muraoka, M. Yasuno, T. Igari, M. Koike, M. Chiba, and T. Mori. 1997. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat. Genet. 17:271-272.
46. Miyakura, Y., K. Sugano, F. Konishi, N. Fukayama, S. Igarashi, K. Kotake, T. Matsui, Y. Koyama, M. Maekawa, and H. Nagai. 2003. Methylation profile of the MLH1 promoter region and their relationship to colorectal carcinogenesis. Genes Chromosomes Cancer. 36:17-25.
47. Miyakura, Y., K. Sugano, F. Konishi, A. Ichikawa, M. Maekawa, K. Shitoh, S. Igarashi, K. Kotake, Y. Koyama, and H. Nagai. 2001. Extensive methylation of hMLH1 promoter region predominates in proximal colon cancer with microsatellite instability. Gastroenterology 121:1300-1309.
48. Nicolaides, N. C., K. W. Kinzler, and B. Vogelstein. 1995. Analysis of the 5′ region of PMS2 reveals heterogeneous transcripts and a novel overlapping gene. Genomics 29:329-334.
49. Palombo, F., I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, and J. Jiricny. 1996. hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. Curr. Biol. 6:1181-1184.
50. Pascal, E., and R. Tjian. 1991. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev. 5:1646-1656.
51. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.
52. Peltomaki, P., H. F. Vasen, et al. 1997. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. Gastroenterology 113:1146-1158.
53. Rajakumar, R. A., S. Thamotharan, R. K. Menon, and S. U. Devaskar. 1998. Sp1 and Sp3 regulate transcriptional activity of the facilitative glucose transporter isoform-3 gene in mammalian neuroblasts and trophoblasts. J. Biol. Chem. 273:27474-27483.
54. Risinger, J. I., A. Umar, J. Boyd, A. Berchuck, T. A. Kunkel, and J. C. Barrett. 1996. Mutation of MSH3 in endometrial cancer and evidence for its functional role in heteroduplex repair. Nat. Genet. 14:102-105.
55. Sakai, T., N. Ohtani, T. L. McGee, P. D. Robbins, and T. P. Dryja. 1991. Oncogenic germ-line mutations in Sp1 and ATF sites in the human retinoblastoma gene. Nature 353:83-86.
56. Shin, K. H., J. H. Shin, J. H. Kim, and J. G. Park. 2002. Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germ-line mutations in promoter of the hMSH2 gene. Cancer Res. 62:38-42.
57. Sia, E. A., R. J. Kokoska, M. Dominska, P. Greenwell, and T. D. Petes. 1997. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol. Cell. Biol. 17:2851-2858.
58. Suske, G. 1999. The Sp-family of transcription factors. Gene 238:291-300.
59. Suske, G. 2000. Transient transfection of Schneider cells in the study of transcription factors. Methods Mol. Biol. 130:175-187.
60. Szadkowski, M., and J. Jiricny. 2002. Identification and functional characterization of the promoter region of the human *MSH6* gene. Genes Chromosomes Cancer 33:36-46.
61. Umar, A., J. I. Risinger, W. E. Glaab, K. R. Tindall, J. C. Barrett, and T. A. Kunkel. 1998. Functional overlap in mismatch repair by human MSH3 and MSH6. Genetics 148:1637-1646.
62. Urban, A., S. Neukirchen, and K. E. Jaeger. 1997. A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. Nucleic Acids Res. 25:2227-2228.
63. Vaisman, A., M. Varchenko, A. Umar, T. A. Kunkel, J. I. Risinger, J. C. Barrett, T. C. Hamilton, and S. G. Chaney. 1998. The role of hMLH1, hMSH3, and hMSH6 defects in cisplatin and oxaliplatin resistance: correlation with replicative bypass of platinum-DNA adducts. Cancer Res. 58:3579-3585.

64. Viel, A., E. Novella, M. Genuardi, E. Capozzi, M. Fornasarig, M. Pedroni, M. Santarosa, M. P. De Leon, L. Della Puppa, M. Anti, and M. Boiocchi. 1998. Lack of PMS2 gene-truncating mutations in patients with hereditary colorectal cancer. Int. J. Oncol. 13:565-569.
65. Wagner, A., Y. Hendriks, E. J. Meijers-Heijboer, W. J. de Leeuw, H. Morreau, R. Hofstra, C. Tops, E. Bik, A. H. Brocker-Vriends, C. van Der Meer, D. Lindhout, H. F. Vasen, M. H. Breuning, C. J. Cornelisse, C. van Krimpen, M. F. Niermeijer, A. H. Zwinderman, J. Wijnen, and R. Fodde. 2001. Atypical HNPCC owing to MSH6 germline mutations: analysis of a large Dutch pedigree. J. Med. Genet. 38:318-322.
66. Wang, Q., C. Lasset, F. Desseigne, J. C. Saurin, C. Maugard, C. Navarro, E. Ruano, L. Descos, V. Trillet-Lenoir, J. F. Bosset, and A. Puisieux. 1999. Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and hMSH6 genes in 75 French kindreds with nonpolyposis colorectal cancer. Hum. Genet. 105:79-85.
67. Wijnen, J., W. de Leeuw, H. Vasen, H. van der Klift, P. Moller, A. Stormorken, H. Meijers-Heijboer, D. Lindhout, F. Menko, S. Vossen, G. Moslein, C. Tops, A. Brocker-Vriends, Y. Wu, R. Hofstra, R. Sijmons, C. Cornelisse, H. Morreau, and R. Fodde. 1999. Familial endometrial cancer in female carriers of MSH6 germline mutations. Nat. Genet. 23:142-144.
68. Wu, Y., M. J. Berends, R. G. Mensink, C. Kempinga, R. H. Sijmons, A. G. van Der Zee, H. Hollema, J. H. Kleibeuker, C. H. Buys, and R. M. Hofstra. 1999. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. Am. J. Hum. Genet. 65:1291-1298.
69. Wu, Y., M. J. Berends, J. G. Post, R. G. Mensink, E. Verlind, T. Van Der Sluis, C. Kempinga, R. H. Sijmons, A. G. van der Zee, H. Hollema, J. H. Kleibeuker, C. H. Buys, and R. M. Hofstra. 2001. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology 120:1580-1587.
70. Wu, Y., M. J. Berends, R. H. Sijmons, R. G. Mensink, E. Verlind, K. A. Kooi, T. van der Sluis, C. Kempinga, A. G. van der Zee, H. Hollema, C. H. Buys, J. H. Kleibeuker, and R. M. Hofstra. 2001. A role for MLH3 in hereditary nonpolyposis colorectal cancer. Nat. Genet. 29:137-138.
71. Yanagisawa, Y., E. Ito, Y. Iwahashi, Y. Akiyama, Y. Yuasa, and K. Maruyama. 1998. Isolation and characterization of the 5′ region of the human mismatch repair gene hPMS1. Biochem. Biophys. Res. Commun. 243:738-743.
72. Zaveri, N. T., C. J. Green, W. E. Polgar, N. Huynh, and L. Toll. 2002. Regulation of transcription of the human prepronociceptin gene by Sp1. Gene 290:45-52.
73. Zhao, L., and L. S. Chang. 1997. The human POLD1 gene: identification of an upstream activator sequence, activation by Sp1 and Sp3, and cell cycle regulation. J. Biol. Chem. 272:4869-4882.

Chapter 7

General discussion

Colorectal tumors arise as a result of a series of genetic changes, starting with single cell lesions and progressing through small benign tumors (adenoma) to malignant cancers (carcinomas). Most colorectal cancers occur in the absence of recognized inherited genetic alterations and are considered sporadic tumors. However, several inherited predispositions to colorectal cancer exist. The two most common and best characterized are hereditary nonpolyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP). Both are dominant transmission syndromes and the inherited predisposition of colorectal cancer results from the inheritance of a single defective mutated gene from the affected parent. HNPCC and FAP account only for a small fraction of colorectal cancers. The study of these inherited forms allows the possibility to discover the responsible genes and to identify the mechanisms involved in colorectal tumorigenesis. A growing amount of genetic and biochemical data indicate that neoplastic transformation in HNPCC and FAP develop by two different molecular pathways (**Chapter 1.1.1**). These are recognized as a series of genetic changes that start with the inactivation of stability genes (mismatch repair [MMR] genes) in HNPCC or a tumor suppressor gene (adenomatous polyposis coli [APC] gene) in FAP. Deficiency of APC protein is thought to result in adenomas and malignant transformation (**Chapter 1.3**). Defects in stability genes do not directly affect cellular growth. The function of the MMR genes is to maintain the integrity of DNA during replication (**Chapter 1.2.1.1**). MMR genes correct nucleotide base mispairs, small insertions and deletions generated by misincorporations or slippage of DNA-polymerase that have escaped proofreading during DNA replication. The MMR genes include: *MLH1*, *MSH2*, *MSH6*, *MLH3*, *PMS1* and *PMS2*. This thesis is focused on the mechanisms that are involved in the first step of the tumorigenesis of HNPCC and on the molecular analysis able to detect genetic alterations in HNPCC and FAP patients and individuals at risk of these hereditary forms of colorectal cancer.

FAP and HNPCC follow the Knudson two-hit model (**Chapter 1.2.1.2**), in which inactivation of the gene arises from an inherited mutated allele at the germline level (first hit) and a second independent genetic alteration on the wild-type allele at the somatic level during the life-time (second hit). Germline mutations result in an inherited susceptibility (high risk) for colorectal cancer in 50% of the offspring of the affected parent. Appropriate positive family history of cancer, detailed histopathology and clinical reports are still crucial to allow mutation analysis. The use of different techniques along with MSI testing and immunohistochemistry has proven useful in detecting MMR defects in suspected HNPCC families [94-96]. In immunohistochemical analysis, a specific genetic deficiency is detected by using a monoclonal antibody against MMR proteins. The inactivation of *MLH1* or *MSH2* genes results in loss of the protein expression in the nuclei of neoplastic cells. When there is a lack of nuclear staining a subsequent mutational analysis is

initially performed on the gene. If a mutation is detected, family members of carriers of the identified genetic alteration are counseled and in those found to be at risk are offered mutation test. Therefore, efforts have been made to find the best clinical criteria based on family history and sensitive screening methods (**Chapter 1.2**), allowing the selection of familial cases with affected members and predisposed individuals. Identification of individuals at high risk of FAP and HNPCC may improve their clinical surveillance and reduce the incidence of colorectal cancer. In the cases where different direct mutational analyses do not succeed to identify genetic alterations, the indirect haplotype analysis is the easiest next step for performing presymptomatic diagnosis (Appendix). In families where several affected individuals are available linkage analysis can be performed, which can be a rapid and still valuable method [1-5]. In **Chapter 2** large set of samples comprising 217 individuals from 20 FAP families has been analyzed by using seven polymorphic markers. Clinical diagnosis of FAP was established in 83 subjects at the moment of the analysis and no mutations were identified in the *APC* gene by classical techniques. The combination of polymorphic markers proved to be informative and allowed an accurate determination of the pedigree structure in 18 out of 20 of the families. Based on the analysis of the affected patient it was possible to identify the haplotype associated with the disease and follow the segregation in all relatives. Ten of the 53 presymptomatic individuals appeared to have inherited the haplotype associated with the affected members and carry the risk to develop FAP. In conclusion, we showed that linkage analysis is still a successful method, providing an opportunity to follow up all 10 asymptomatic individuals with clinical screening tests and allowing the detection of the disease at an early age and curable stage.

Linkage analysis is less often used in diseases such as HNPCC, where heterogeneity of genetic alterations and multiple genes can be involved (**Chapter 1.2.3**). In HNPCC cases the first test performed, as recommended by ICG-HNPCC [6], should be Microsatellite Instability (MSI) analysis (**Chapter 1.2.2.1**). Instable microsatellites have been considered to reflect mutant MMR genes. This instability occurs in somatic cells (tumor tissue) as a consequence of inactivation of the MMR system and is identified as small insertions or deletions within short tandem repeats (microsatellite repeats) located throughout the genome (**Chapter 1.2.2**). According to international guidelines [6], MSI phenotype is classified in two distinct groups of tumors: MSI-H (high frequency of MSI) and MSI-L (low frequency of MSI). The use of greater a number of polymorphic markers allows the distinction between MSI-L and MSS (microsatellite stable) (**Chapter 3**). MSI analysis on 45 colorectal cancer patients, including 13 HNPCC families and 32 non-HNPCC families, showed MSI-H status in 85% of HNPCC families and in 16% of non-HNPCC families (**Chapter 3**). We observed an association between HNPCC and MSI-H, as also reported by others[7-9], providing

support for the hypothesis [10] that a "mutator" phenotype may promote tumorigenesis. Moreover, MSI-H status has also been correlated to increased risk of cancer in the proximal colon, where HNPCC tumors are typically localized. On the basis of these conclusions, MSI analysis appears to be a convenient and rapid prescreening test for HNPCC families fulfilling the Amsterdam (I-II) and Bethesda criteria.

HNPCC cases with MSI-H have been correlated with the presence of genetic alterations in the MMR genes, in particular *MLH1* or *MSH2* [11-14]. In agreement with this finding, 50% of affected HNPCC patients showing MSI-H tumors have been identified to carry a germline mutation in one of these two genes (**Chapter 4**). The mean age of diagnosis of these patients was significantly lower (about 10 years) compared to patients without mutations, suggesting that mutations in the *MLH1* or *MSH2* genes cause a more severe phenotype with an early age of tumor onset. In these patients no genetic alterations were detected in the *MSH6* gene. The analysis was performed on 16 Italian families by SSCP and direct sequence analysis. To facilitate detection of mutations in the *MSH2* and *MSH6* genes a protein truncation test (PTT) was added. The choice of these methods was based on the type of genetic alterations previously identified in the MMR genes, which have been assembled in a database [15]. The finding in Chapter 3 and 4 establish a consistent correlation between MSI-H and identified germline mutations in HNPCC cases. In addition the results assess the advantage of MSI analysis as the first test with the purpose to identify germline mutation carriers.

Alternative mechanisms of MMR gene inactivation include genomic deletions, promoter methylation and alteration in the binding of transcription factors, where the latter two result in transcriptional silencing of the gene (**Chapter 1.2.1.2**). In humans epigenetic alterations are characterized by methylation of the cytosine nucleotide, giving a 5-methylcytosine at the CpG site with no changes in DNA sequence (**Chapter 1.2.3.4**). Most of the promoters of human genes are rich in CpG sequences (CpG islands) and are therefore susceptible to methylation (an epigenetic process). Promoter methylation as an epigenetic alteration associated with transcriptional silencing of tumor suppressor genes has been reported in different sporadic cancers [16, 17]. In many sporadic colon cancers somatic silencing of both copies of the *MLH1* gene is associated with biallelic methylation, or methylation and LOH, or mutation and promoter methylation, which all result in a loss of MLH1 expression [18-21]. Silencing of the *MLH1* gene by hypermethylation of its 5'CpG islands has been found in most sporadic primary colorectal cancer cases with MSI [22]. The idea that hypermethylation and the subsequent inactivation of the *MLH1* gene leads to MSI is supported by

the observation that reversing the methylation status, by treating the cultured colorectal cancer cells with a demethylating agent (5'-aza-2'deoxycytidine) corrects the MMR defect and the associated MSI [22]. Promoter methylation of the *MLH1* gene has been observed in the majority of colorectal adenocarcinomas [22-25] but also in normal colonic epithelium [26]. Recently, epigenetic alterations in the *MLH1* promoter have been detected in peripheral blood (**Chapter 5**, [27-29]), indicating a new mechanism of allelic inactivation at the germline level (**Chapter 1.2.1.2**). In **Chapter 5** we report the first case of constitutional hemi-methylation in the promoter of the *MLH1* gene. Fourteen HNPCC cases with MSI-H and without germline mutations in the MMR genes were examined for hypermethylation of the *MLH1* gene using a methylation-specific polymerase (MSP) chain reaction assay (**Chapter 1.2.3.4**). The analyzed promoter region was previously reported as a critical area associated with a loss of expression of the gene due to the possible presence of several methylated CpG sites close to the regulatory motif and transcription factor binding sites [30]. One of the 14 HNPCC cases was identified to have methylated and unmethylated alleles in blood DNA and only methylated allele in tumor tissue. Analysis of SNP previously described in this region [31] showed that blood DNA was heterozygous for the -93 G/A polymorphism, whereas tumor DNA was homozygous (G), indicating loss of heterozygosity (LOH) and resulting in the loss of the unmethylated allele at somatic level. Consistent with a loss of the unmethylated allele immunohistochemistry showed an absence of MLH1 protein in the tumor. This single case raised the possibility that methylation at the germline level represents an alternative genetic alteration to the well-known missense, nonsense mutations and genomic deletions. In the following years other investigators [27-29] have reported cases of constitutional hemi-methylation in HNPCC and sporadic colorectal cancer at an early age with MSI-H using MSP assays. Suter et al. [27] identified two unrelated individuals with hypermethylation in the promoter of the MLH1 gene. Both patients were characterized as MSI-H HNPCC cases with a clinical diagnosis of the first colorectal cancers at the age of 43 and 46. Epigenetic alterations were observed in all normal tissues, including peripheral blood, buccal mucosa and hair follicles. *MLH1* methylation was also detected at a low level in germ cells (spermatozoa). In both HNPCC patients methylation of the *MLH1* promoter was confined to one of the two alleles and the unmethylated allele was lost in the tumors, LOH was detected using a single nucleotide polymorphic site as previously mentioned ([31] and **Chapter 5**). A complete loss of MLH1 protein expression was observed using immunohistochemistry. Using the same technique, germline *MLH1* epimutation was also identified in a single HNPCC patient by Hitchins et al. [29] Hypermethylation of the *MLH1* promoter was detected in peripheral blood in 4 out of 30 sporadic early-onset colon cancer patients with MSI-H [28]. In three cases methylation was also identified in both normal tissue (colon, gastric mucosa, endometrium and bone marrow) and tumor tissue (colon

and endometrial cancer). Analysis of the polymorphic site in the MLH1 promoter showed hemiallelic methylation in two of these three cases. In one case the unmethylated allele was lost in the tumor. Interestingly, all above cases follow the two-hit tumor suppressor gene hypothesis (**Chapter 1.2.1.2**), confirming that biallelic inactivation by promoter methylation and LOH result in an absence of MLH1 expression in the tumor. Recently, a first case of germline hypermethylation and somatic frameshift mutation of the *MSH2* gene has been reported by Chan et al. Methylation of *MSH2* promoter was observed in one of the 31 analyzed cases with early-onset or familial CRCs and MSI. Subsequent methylation analysis was performed in the epimutation carrier's family and ten of the twelve family members were identified to have hypermethylation of *MSH2* promoter in blood leukocytes and in normal rectal mucosa. Analysis of two SNPs in the promoter region showed that hypermethylation was always associated with two specific variants (-118T;-443G) in blood and in somatic tissues. Haplotype analysis showed a perfect segregation of the haplotype associated with cancer (2 CRC; 1 EndoCA) in ten members of the family, all carrying germline *MSH2* methylation. In three siblings, cancer was detected at early age (CRC-30yrs; EndoCA-43yrs; CRC-40yrs) and tumors did not show expression of MSH2 protein. Only two individuals did not have inherited the disease haplotype and methylation was also not observed in blood leukocytes, rectal mucosa or buccal mucosa. Methylation analysis was also performed in 56 paired tumor and normal sporadic CRC samples, 11 colorectal adenomas and 10 germline DNA samples from blood donors, but none of them showed hypermethylation in *MSH2* promoter. Chan et al. have reported the first case of inherited epimutation in a family putatively affected with HNPCC, suggesting that hypermethylation could act as the first "hit" to inactivate one allele of *MSH2*, whereas second hit is acquired in the hemimethylated cells for the progression of cancer. In conclusion, these data confirm our earlier findings in Chapter 5 and further proof that methylation and the associated silencing of the *MLH1*and *MSH2* genes represent an alternative germline change that can underlie the development of cancer.

Another mechanism has been proposed for constitutional allelic inactivation involving a change at the gene expression level rather than alteration in the DNA sequence. Different investigators (**Chapter 6**, [32-36]) have reported that single nucleotide polymorphisms (SNPs) in the regulatory promoter region are associated with altered promoter activity and gene expression, both *in vivo* and *in vitro*. In genes such as *Rb*, *EGFR* and *hMSH6* [33, 37, 38], whose promoter activity is mostly regulated by Sp1 transcription factors in the absence of a TATA box, the presence of SNPs at the Sp1 sites can quantitatively alter mRNA levels and protein expression resulting in a susceptibility for tumorigenesis. Functional analysis of the 5'-flanking region of the hMSH6 gene identified two

single nucleotide polymorphisms located within the consensus sequence of two of the seven Sp1 transcription factor binding sites in the core promoter (**Chapter 6**). Polymorphism analysis revealed that the two SNPs completely eliminated the ability to bind Sp1 protein, resulting in a 50% reduction in promoter activity compared to the wild-type allele. This finding was confirmed by *in vivo* experiments, in which *MSH6* mRNA and protein were found at very low levels in glioblastoma cell lines containing both SNPs in heterozygous or homozygous form. Previously analysis of identified SNPs in the promoter region has never been observed to affect the transcription of the MMR genes [31, 39, 40, 41]. A subsequent analysis on Caucasian and diversity sample sets showed that the two allelic variants were always associated with a higher incidence in the Caucasian population (16% vs. 3%). Mutational and biochemical analyses were performed to characterize the regulation of the *MSH6* promoter and the function of the seven potential Sp1 binding sites. We observed that all seven Sp1 binding sites contributed to the expression of the MSH6 gene, but the four Sp1 sites closest to the transcriptional start sites were the most important. Synergistic activity, due to binding of multiple Sp1 transcription factors, has been previously reported [42, 43, 44] in other genes. In addition, all Sp1 sites were found to bind Sp1 and Sp3 transcription factors. We also observed that competition for these sites and the promoter activity depended on the intracellular levels of the transcriptional activator (Sp1) and the repressor (Sp3) factor. This finding is consistent with the hypothetical model of Kwon et al. [45], which suggests that the occupied number of Sp1 and Sp3 sites are determined by the proportion of the relative transcription factors in the cell. In addition, using an *in vitro* assay we observed that methylation of the seven Sp1 sites decreased the ability of each site to bind Sp1 and Sp3 resulting in a 90% reduction of *MSH6* promoter activity. Methylation has been found to inhibit Sp1 binding sites depending on the position of the methylated C in the Sp1 site (CCGCCC) [46]. On the other hand, binding of the Sp1 transcription factor to the GGGCGG/CCGCCC box also helps to maintain the CpG islands in an unmethylated state [46]. *In vitro* methylation analysis of the two Sp1 sites containing the SNPs showed a decrease of the promoter activity to <1%, suggesting that genetic alterations at the Sp1 site could be the target of de novo methylation and be more susceptible to silencing by epigenetic process compared to the wild-type allele with seven Sp1 sites. In conclusion, these data showed that polymorphisms in the regulatory region increase a decreased constitutional expression of one allele of the *MSH6* gene. Moreover, the results suggest that the human population contains individuals with different levels of MHS6 promoter activity. A possible consequence consistent with the Knudson two-hit model is that additional mutations and loss of the wild type allele at the somatic level can result in inactivation of the *MSH6* gene. Silencing of this gene causes an increased accumulation of base substitution mutations, single base insertion/deletion

mutations and instability of the mononucleotide repeats (**Chapter 1.2.3.2**). Absence of *MSH6* protein results in a partial loss of mismatch repair function, due to the partial redundancy function with *MSH3* gene, which repairs large insertion/deletion mutations (**Chapter 1.2.1.1**).

In conclusion, we identified different mechanisms that can lead to the inactivation of MMR genes in HNPCC, including a high incidence of point mutations. Most interestingly, two novel mechanisms were identified that occurred at a low incidence: (i) methylation of the *MLH1* gene, and (ii) SNPs in the transcription factor sites of the *MSH6* gene. Both genetic alterations affected the expression of one of the alleles at a constitutional level. Recently other groups have identified new cases of epigenetic alteration of *MLH1* in mutation negative carriers. Thus, additional analyses may improve the detection of genetic alterations in mutation negative individuals, such as an analysis of the transcriptional regulation of the MMR promoter regions and the methylation status of *MLH1* gene. In a large fraction of HNPCC and FAP patients advanced screening techniques do not detect any genetic alterations. In these cases we showed that linkage analysis may help identify the individuals at risk of hereditary colorectal cancer.

References

1. Dunlop MG, Wyllie AH, Steel CM, Piris J, Evans HJ. Linked DNA markers for presymptomatic diagnosis of familial adenomatous polyposis. Lancet 1991;337:313-6.
2. Mareni C, Stella A, Origone P, Susca F, Montera MP, Lonoce A, Ponz de Leon M, Sassatelli R, Gentile M, Straface A, et al. Linkage studies in Italian families with familial adenomatous polyposis. Hum Genet 1993;90:545-50.
3. Park JG, Han HJ, Kang MS, Nakamura Y. Presymptomatic diagnosis of familial adenomatous polyposis coli. Dis Colon Rectum 1994;37:700-7.
4. Walpole IR, Kool DA, Edkins T, Creegan R, Levitt S, Francis ST, Goldblatt J. Genetic counselling and gene mutation analysis in familial adenomatous polyposis in Western Australia. Med J Aust 1995;162:464-7.
5. Amikam D, Niv D, Lachter J, Eidelman S, Ben-Ishai Z. The use of DNA markers in the pre-clinical diagnosis of familial adenomatous polyposis. Isr J Med Sci 1997;33:8-13.
6. 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.
7. Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Peltomaki P, de la Chapelle A, Hamilton SR, Vogelstein B, Kinzler KW. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat Med 1996;2:169-74.
8. Giardiello FM, Brensinger JD, Petersen GM. AGA technical review on hereditary colorectal cancer and genetic testing. Gastroenterology 2001;121:198-213.
9. Aaltonen LA, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomaki P, Chadwick RB, Kaariainen H, Eskelinen M, Jarvinen H, Mecklin JP, de la Chapelle A. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 1998;338:1481-7.
10. Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res 1991;51:3075-9.

11. Thibodeau SN, French AJ, Roche PC, Cunningham JM, Tester DJ, Lindor NM, Moslein G, Baker SM, Liskay RM, Burgart LJ, Honchel R, Halling KC. Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes. Cancer Res 1996;56:4836-40.
12. Olschwang S, Hamelin R, Laurent-Puig P, Thuille B, De Rycke Y, 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.
13. Wahlberg SS, Schmeits J, Thomas G, Loda M, Garber J, Syngal S, Kolodner RD, Fox E. Evaluation of microsatellite instability and immunohistochemistry for the prediction of germ-line MSH2 and MLH1 mutations in hereditary nonpolyposis colon cancer families. Cancer Res 2002;62:3485-92.
14. Lynch HT, Lynch JF. Lynch syndrome: history and current status. Dis Markers 2004;20:181-98.
15. Peltomaki P, Vasen HF. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology 1997;113:1146-58.
16. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999;21:163-7.
17. Rountree MR, Bachman KE, Herman JG, Baylin SB. DNA methylation, chromatin inheritance, and cancer. Oncogene 2001;20:3156-65.
18. 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.
19. Kuismanen SA, Holmberg MT, Salovaara R, de la Chapelle A, 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.
20. Furukawa T, Konishi F, Masubuchi S, Shitoh K, Nagai H, Tsukamoto T. Densely methylated MLH1 promoter correlates with decreased mRNA expression in sporadic colorectal cancers. Genes Chromosomes Cancer 2002;35:1-10.
21. Menigatti M, Di Gregorio C, Borghi F, Sala E, Scarselli A, Pedroni M, Foroni M, Benatti P, Roncucci L, Ponz de Leon M, Percesepe A. Methylation pattern of different regions of the MLH1 promoter and silencing of gene expression in hereditary and sporadic colorectal cancer. Genes Chromosomes Cancer 2001;31:357-61.
22. 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.
23. 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.
24. Esteller M, Levine R, Baylin SB, Ellenson LH, Herman JG. MLH1 promoter hypermethylation is associated with the microsatellite instability phenotype in sporadic endometrial carcinomas. Oncogene 1998;17:2413-7.
25. Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD, Markowitz SD. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95:8698-702.
26. Nakagawa H, Nuovo GJ, Zervos EE, Martin EW, Jr., Salovaara R, Aaltonen LA, de la Chapelle A. Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. Cancer Res 2001;61:6991-5.
27. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004;36:497-501.
28. Miyakura Y, Sugano K, Akasu T, Yoshida T, Maekawa M, Saitoh S, Sasaki H, Nomizu T, Konishi F, Fujita S, Moriya Y, Nagai H. Extensive but hemiallelic methylation of the hMLH1 promoter region in early-onset sporadic colon cancers with microsatellite instability. Clin Gastroenterol Hepatol 2004;2:147-56.
29. Hitchins M, Williams R, Cheong K, Halani N, Lin VA, Packham D, Ku S, Buckle A, Hawkins N, Burn J, Gallinger S, Goldblatt J, Kirk J, Tomlinson I, Scott R, Spigelman A, Suter C, Martin D, Suthers G, Ward R. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;129:1392-9.
30. 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;59:2029-33.
31. Ito E, Yanagisawa Y, Iwahashi Y, Suzuki Y, Nagasaki H, Akiyama Y, Sugano S, Yuasa Y, Maruyama K. A core promoter and a frequent single-nucleotide polymorphism of the mismatch repair gene hMLH1. Biochem Biophys Res Commun 1999;256:488-94.
32. Baseggio L, Bartholin L, Chantome A, Charlot C, Rimokh R, Salles G. Allele-specific binding to the -308 single nucleotide polymorphism site in the tumour necrosis factor-alpha promoter. Eur J Immunogenet 2004;31:15-9.

33. Gazzoli I, Kolodner RD. Regulation of the human MSH6 gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms. Mol Cell Biol 2003;23:7992-8007.
34. Harendza S, Lovett DH, Panzer U, Lukacs Z, Kuhnl P, Stahl RA. Linked common polymorphisms in the gelatinase a promoter are associated with diminished transcriptional response to estrogen and genetic fitness. J Biol Chem 2003;278:20490-9.
35. Mann V, Hobson EE, Li B, Stewart TL, Grant SF, Robins SP, Aspden RM, Ralston SH. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. J Clin Invest 2001;107:899-907.
36. Stevens A, Soden J, Brenchley PE, Ralph S, Ray DW. Haplotype analysis of the polymorphic human vascular endothelial growth factor gene promoter. Cancer Res 2003;63:812-6.
37. Sakai T, Ohtani N, McGee TL, Robbins PD, Dryja TP. Oncogenic germ-line mutations in Sp1 and ATF sites in the human retinoblastoma gene. Nature 1991;353:83-6.
38. Liu W, Innocenti F, Wu MH, Desai AA, Dolan ME, Cook EH, Jr., Ratain MJ. A functional common polymorphism in a Sp1 recognition site of the epidermal growth factor receptor gene promoter. Cancer Res 2005;65:46-53.
39. Iwahashi Y, Ito E, Yanagisawa Y, Akiyama Y, Yuasa Y, Onodera T, Maruyama K. Promoter analysis of the human mismatch repair gene hMSH2. Gene 1998;213:141-7.
40. Shin KH, Shin JH, Kim JH, Park JG. Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germ-line mutations in promoter of the hMSH2 gene. Cancer Res 2002;62:38-42.
41. Yanagisawa Y, Ito E, Iwahashi Y, Akiyama Y, Yuasa Y, Maruyama K. Isolation and characterization of the 5' region of the human mismatch repair gene hPMS1. Biochem Biophys Res Commun 1998;243:738-43.
42. Hagen G, Muller S, Beato M, Suske G. Sp1-mediated transcriptional activation is repressed by Sp3. Embo J 1994;13:3843-51.
43. Mastrangelo IA, Courey AJ, Wall JS, Jackson SP, Hough PV. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. Proc Natl Acad Sci U S A 1991;88:5670-4.
44. Pascal E, Tjian R. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev 1991;5:1646-56.
45. Kwon HS, Kim MS, Edenberg HJ, Hur MW. Sp3 and Sp4 can repress transcription by competing with Sp1 for the core cis-elements on the human ADH5/FDH minimal promoter. J Biol Chem 1999;274:20-8.
46. Clark SJ, Harrison J, Molloy PL. Sp1 binding is inhibited by (m)Cp(m)CpG methylation. Gene 1997;195:67-71.

Appendix

Screening strategies and approaches

This section focuses on the types of strategies used in laboratories conducting screening analysis and describes the benefits and disadvantages of the different techniques. Some of the assays discussed below have been used in this study and are described in the following Chapter. A particular emphasis will be put on successful novel techniques. Automated and high-throughput analytical technology has allowed to switch from the use of classical radioactive molecules to the fluorescent molecules, resulting in sensitive assays, no radioactive score and a lower risk for the operator.

1. Classical methods

Predictive genetic testing is indispensable for individuals at risk of colorectal cancer in order to receive an appropriate surveillance program in a timely manner. The detection rate can be improved by using multiple techniques that reliably detect point mutations, insertions, rearrangements, hypermethylation, deletions of individual base pairs, entire genes, or whole chromosomes, with a cumulative sensitivity of over 95%. Linkage analysis and direct mutational analysis are the two main approaches used in genetic testing. Linkage analysis, which is used in **Chapter 2**, compares the haplotypes of affected patients with unaffected members of the family. The haplotype is found by analyzing di-trinucleotides repeats (genetic markers) located within or close to the chromosomal region of the gene. Linkage analysis has been found useful in the identification of new genes by localization of loss of heterozygosity or as a first step in the presymptomatic diagnosis of monogenic diseases, which are due to a mutation in a single locus. Such analysis requires DNA from the affected and unaffected individuals, where a small number of samples, uncertain paternity or early death can restrict the usefulness. Moreover, linkage analysis provides only an estimate of the risk of being a carrier of the mutant gene and therefore laboratories do not perform this method anymore. In contrast, direct mutational analysis offers the possibility to detect precise location of the mutation. The choice of the techniques is based on the gene involved, the sensitivity of the assay and the cost-effectiveness of the test. The existing tests fall in two categories: (1) screening techniques, which are able to detect mutations that result in conformational changes of the DNA or expression of the gene, and (2) direct sequencing of DNA. In diseases such as HNPCC where different genes can be involved direct sequencing analysis can be very expensive and in many laboratories it is performed only after classical or novel methods have detected a mutation. Direct sequencing analysis has some limitations in that it can not detect large genomic deletions or mutations that affect gene expression, and the sensitivity of the assay is 80% [1].

For HNPCC cases the ICG-HNPCC recommends immunohistochemistry (IHC) and microsatellite instability analysis (MSI) before undertaking any other screening methods. The majority of FAP cases are analyzed using a protein truncation test (PTT). The use of this technique is explicable because most mutations are insertions, deletions and nonsense mutations that lead to frame shifts or premature stop codons, which result in the truncation of the APC gene product. Protein from APC genes with a truncating mutation is often unstable and difficult to detect by Western blotting [2]. The PTT assay uses in vitro transcription and translation to generate a protein product from mRNA, which results in a shorter size product compared to the wild-type allele if it carries a mutation such as a deletion, insertion or splice variant. Unfortunately, it does not detect missense mutations, in-frame deletions, insertions or promoter mutations. Moreover, a potential problem using cDNA is that the level of mutant mRNA could be much lower than that of the wild-type mRNA. This altered allele ratio can be detected using an allele-specific-expression assay which in combination with PTT analysis increases the detection of APC gene alterations to 87% [3]. Alternative strategies are highly sensitive and specific for detecting of changes in DNA structure created by point mutations. These classical techniques include single-strand conformation polymorphisms (SSCP), denaturing gradient gel electrophoresis (DGGE) and denaturing high-performance liquid chromatography (DHPLC) assays. The latter two techniques are the most popular heteroduplex assays and are used as a first screen to identify point mutation, small deletions and duplications. DGGE is an alternative to PTT and has a higher sensitivity of detecting mutations in the APC gene [4]. DGGE requires expensive special primers with GC clamp, the choice of which is very critical for the result of the assay. Different authors indicate the combined use of DGGE and PTT, or Southern blot and direct sequencing as an efficient strategy for the detection of mutations in FAP [5-7]. The use of DHPLC for the detection of DNA sequence variations in APC and MMR genes has been found to be a rapid and efficient method with a sensitivity of 97% [8-11]. The method is based on a comparative analysis of the mobility of the heteroduplex and the wild-type homoduplex under partially denaturing conditions using different melting temperatures and automated HPLC equipped with a DNA separation column. The optimum conditions are determined by the sample size and the melting behavior of the wild-type sequence, resulting in an early elution of the sample due to a decreased retention time by the DNA separation column. Normally, the heteroduplex results in a reduced column retention time compared to the homoduplex strand. Heteroduplexes are confirmed by sequence analysis, determining the precise location of the mutations. This method presents several advantages including the use of DNA and amplified products, a rapid result (7 minutes for sample), a high sensitivity, reproducible data and low cost compared to others. However, like DGGE and direct sequence analysis, DHPLC is limited by the inability to detect gross deletions. In the last two

decades the SSCP assay has been the most common and cheapest analysis, although it is less sensitive to detect mutations and DNA polymorphisms. SSCP is based on the ability of one or more changed nucleotides to alter the electrophoretic mobility pattern of a single-stranded DNA molecule under non-denaturing conditions (**Chapter 3**).

2. Fluorescent techniques

Radioisotopic and non-radioisotopic techniques have been replaced by fluorescent detection methods. For example, in the SSCP assay detection with radioisotope or silver staining (**Chapter 3**) has been substituted by fluorescent SSCP (F-SSCP). The new technique is based on the use of different fluorescent dyes containing chemical group (fluorophores) that fluoresce when exposed to a specific wavelength of light. Fluorescent dyes are used as labels and can be incorporated directly during the polymerase chain reaction (PCR) by being attached to a nucleic acid (ddNTPs), or indirectly by being attached to a primer. The fluorescent dyes are excited by a beam of light from an appropriate light source (argon laser) through a color filter (excitation filter) designed to transmit light at the desired excitation wavelength. During the electrophoresis run the laser beam is focused at a specific constant position on the gel, where it excites the migrated fluorophore attached to the sample resulting in emission of light that is captured by a barrier filter and transmitted to a charged coupled device (CCD) camera. It is possible to combine multiple fluorophores in the same reaction with different emission and excitation wavelengths. In the F-SSCP assay two fluorescent dyes are attached to the primers and the amplification products result in two peaks of different mobility in the presence of the heteroduplex [12]. In the fluorescent competitive allele-specific polymerase chain reaction (SNiPTag) [13] assay, fluorescence is only detected when the extreme 3' terminus of the primers match perfectly with the template and permit the final synthesis of the strands. However, in direct sequencing the use of four different fluorescent dyes labeled to the four different chain-terminating dideoxynucleotides (G, A, T, C) is necessary to determine the complete sequence of the fragment. In other methods such as multiplex ligation-dependent probe amplification (MLPA) and multiplex amplification and probe hybridization (MAPH) [14], one fluorescent dye is used to detect multiplex amplification products by a difference in amplicon sizes. In the last few years MLPA and MAPH have detected large numbers of genomic rearrangements that had not been revealed by conventional DNA and RNA-based techniques [15-19]. Both methods have proven to be efficient and reliable in identifying deletions and duplications in several genes, including MMR and APC genes. MLPA is a quantitative multiplex PCR-based method, which allows the identification up to 40

different PCR products of the DNA target sequence gene in a single experiment, requiring a small amount of genomic template [20]. Specific probes for each exon consist of two oligonucleotides (hemiprobe) containing target-specific sequences at one site and a universal primer annealing sequence at the other site (Fig. 1).

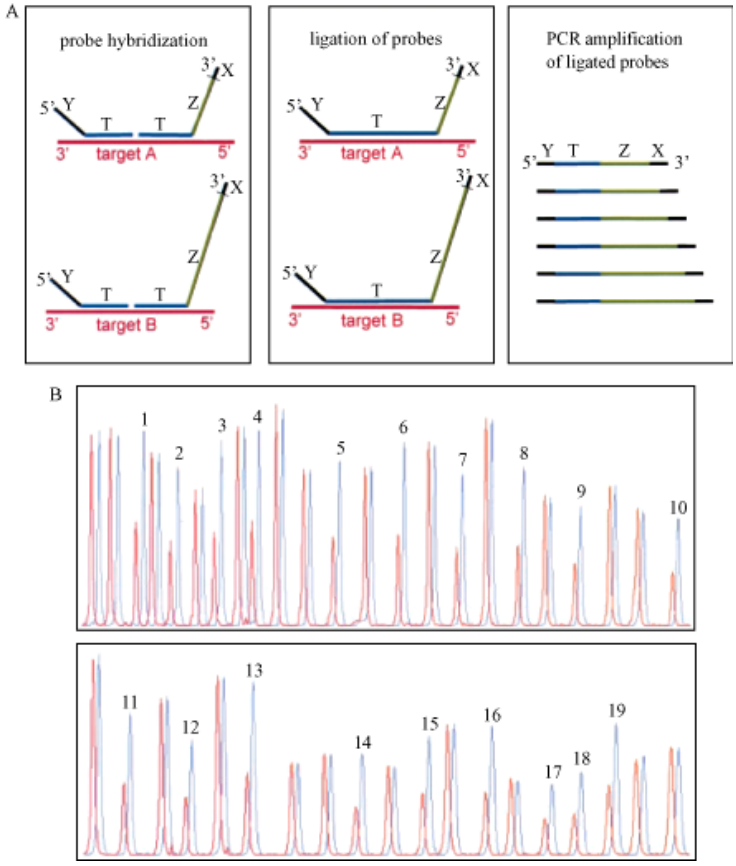

Figure 1. Graphic example of Multiplex Ligation-dependent Probe Amplification (MLPA) technique. **A**. Denatured genomic DNA is hybridized with a mixture probes (T). Each probe consists of two oligonucleotides that hybridize to the adjacent target sequence (target A or B), two universal primer (X, Y) and different stuffer sequence (Z). The two hemi-probes are ligated to produce a unique probe (T) by a thermostable ligase. All probe ligation products are amplified simultaneously by PCR using universal primers (X and Y), resulting in unique amplification length. **B**. Amplification products are separated by capillary electrophoresis and relative amounts of probe amplification products reflect the relative copy number of target sequences. An example of deletion of the entire MLH1 gene is reported. The exons are labelled from 1 to 19 and each one is characterized by one short peak showing deletion and one tall peak representing the control DNA. Unlabelled peaks are internal control (MSH2 exons and other genes). Adapted from Gille JJ et al., Br J Cancer. 2002 Oct 7; 87(8):892-7. Reproduced with permission from Nature Publishing Group.

Both hemiprobe are designed to hybridize to adjacent target sequences. After hybridization both hemiprobes are ligated to form one strand permitting the subsequent amplification. In the presence of genomic rearrangements the two oligonucleotides are not located next to each other resulting in an absence of ligation and amplification. The ligated probes for different exons or for different genes have an identical end sequence (universal primer) and a different length of the stuffer sequence, which allows the simultaneous amplification in a single PCR reaction and electrophoretic separation by size gradient. The amount of PCR product is proportional to the amount of the specific target sequence present in the sample. Semiquantitative analysis is performed by attaching a fluorophore to the primers and comparing the peak heights of each amplicons to the corresponding exons in the normal control. The evaluation is carried out by the ratio between the interesting peak and its wild-type, where 1.0 is the dosage quotient for normal sequence, 0.5 (range 0.3-0.7) and 1.5 (range 1.3-1.7) for deletions and duplications, respectively. MAPH provides an alternative DNA approach to MLPA for detecting and quantifying genomic rearrangements. Like the previous technique this method is based on the hybridization of amplifiable probes to unique genomic DNA sequences. One microgram of DNA is immobilized on a nylon filter (Fig. 2). After hybridization with the probe mixture all nonspecific probes are removed by stringent washing. Hybridized probes are denatured and amplified quantitatively by PCR after which fluorescently labeled products are analyzed by electrophoresis. MLPA and MAPH have improved the sensitivity of mutational analysis in those cases where mutations can not be identified by sequence analysis or classical methods. Both assays have proved to be useful and sensitive methods for the gene dosage analysis of MMR and APC genes [21-25]. They are PCR-based assays and generate similar results. Based on the type of probes MAPH has the advantage of being not sensitive to single substitution variants within the target sequence, since the probe (100-200 base pairs) hybridizes even in the presence of mutations or deletions/insertions of a single base pair in the tested exon. The disadvantage is that MAPH probes are always amplifiable and they need to be removed from the unbound target template, which step results in risk of contamination. The advantage of using MLPA is that the probes become amplifiable after the ligation and the entire reaction is performed in a single tube. Due to the short length of the specific probes (20-30 nucleotides) polymorphisms or single base mutations in the probe binding region can affect probe hybridization, which may result in an absence of ligation and amplification, and an apparent deletion of the exon. Therefore, detection of single exon deletion events should be confirmed either by sequencing long-range PCR product (from DNA or cDNA) or by including two different sets of probes in each exon. The advantage of both techniques is the small amount of DNA required (1 μg for MAPH analysis and only 100-200 ng for MLPA) compared to classical methods, such as Southern blot (2.5 μg).

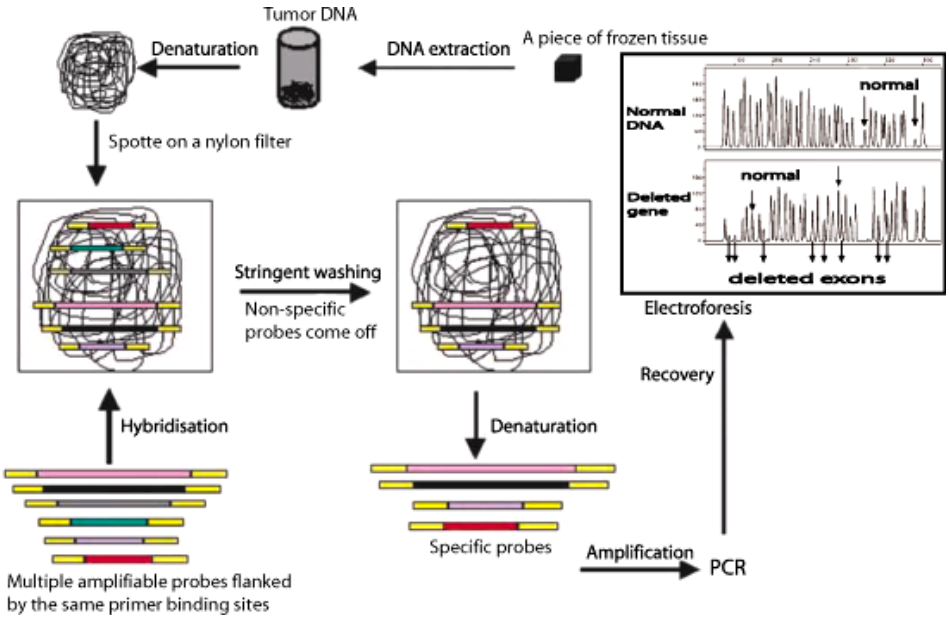

Figure 2. The MAPH technique. Amplificable probes recognize unique regions in the genome and are flanked by the same primer pair. After stringent washing, the specific hybridized probes are denatured, amplified quantitatively by PCR and then analyzed by electrophoresis. Adapted from Rakha E.A. et al., Int. J. Cancer 2005; 114: 720-729. Reproduced with permission from John Wiley & Sons, Inc.

Until recently, Southern blot has been the most popular and most sensitive method used to identify and characterize deletions and duplications in different genes [5, 26, 27]. In this assay genomic DNA is digested with different restriction endonucleases, the resulting fragments of DNA are size-fractionated by gel and transferred to a hybridization membrane where radioactive probes recognize specific target sequences. Excess of unbound radioactive probes are eliminated by washing steps and complexes DNA-probe are detected by X-ray film. The probe contains multiple noncontiguous exons and deletion of one or more exons results in an extra band of shorter size compared to the wild-type allele of the same sample. A major advantage of this method is that multiple-exon probes can directly identify the exons involved in the genomic rearrangement. However substantial amounts of DNA and several steps are required. The estimated working time is around 6 days to complete the analysis (versus 2 days for MLPA). Baudhuin et al. [28] analyzed a set of samples by MLPA and Southern blot, and cross-validated the two methods for the detected genomic deletions. Results from both techniques showed 98.8% concordance, minor discrepancies were represented by the inability of MLPA to detect deletions in the promoter region (which could be resolved by adding probes to cover this area). According to these and other data [17, 21, 25, 28], MLPA is an easy,

sensitive and highly reproducible method that can be included as a standard analysis method in diagnostic laboratories. It is absolutely necessary to determine precisely the breakpoints of all detected deletions and in the cases where the sequence of the rearrangement is localized, a simple diagnostic PCR allows the screening in at-risk family members.

The development of fluorescence technology has improved the identification of genetic alterations and the sensitivity of the analysis. It is applicable to different techniques and offers the possibility to analyze a large set of samples in a short amount of time, to result in simple data and eliminate the use of radioactivity. Different authors indicate the combined use of some of the above techniques and direct sequencing as an efficient strategy for the detection of mutations in FAP and HNPCC.

Advance technology increases the sensitivity of genetic testing, resulting in a decrease of mortality due to a better treatment and follow up in patients carrying genetic alterations.

References

1. Farrington SM, Lin-Goerke J, Ling J, Wang Y, Burczak JD, Robbins DJ, Dunlop MG. Systematic analysis of hMSH2 and hMLH1 in young colon cancer patients and controls. Am J Hum Genet 1998;63:749-59.
2. Smith KJ, Johnson KA, Bryan TM, Hill DE, Markowitz S, Willson JK, Paraskeva C, Petersen GM, Hamilton SR, Vogelstein B, et al. The APC gene product in normal and tumor cells. Proc Natl Acad Sci U S A 1993;90:2846-50.
3. Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, Hamilton SR, Vogelstein B, Kinzler KW. Molecular diagnosis of familial adenomatous polyposis. N Engl J Med 1993;329:1982-7.
4. Fodde R, van der Luijt R, Wijnen J, Tops C, van der Klift H, van Leeuwen-Cornelisse I, Griffioen G, Vasen H, Khan PM. Eight novel inactivating germ line mutations at the APC gene identified by denaturing gradient gel electrophoresis. Genomics 1992;13:1162-8.
5. van der Luijt RB, Khan PM, Vasen HF, Tops CM, van Leeuwen-Cornelisse IS, Wijnen JT, van der Klift HM, Plug RJ, Griffioen G, Fodde R. Molecular analysis of the APC gene in 105 Dutch kindreds with familial adenomatous polyposis: 67 germline mutations identified by DGGE, PTT, and southern analysis. Hum Mutat 1997;9:7-16.
6. Kohoutova M, Stekrova J, Jirasek V, Kapras J. APC germline mutations identified in Czech patients with familial adenomatous polyposis. Hum Mutat 2002;19:460-1.
7. Vandrovcova J, Stekrova J, Kebrdlova V, Kohoutova M. Molecular analysis of the APC and MYH genes in Czech families affected by FAP or multiple adenomas: 13 novel mutations. Hum Mutat 2004;23:397.
8. Holinski-Feder E, Muller-Koch Y, Friedl W, Moeslein G, Keller G, Plaschke J, Ballhausen W, Gross M, Baldwin-Jedele K, Jungck M, Mangold E, Vogelsang H, Schackert HK, Lohsea P, Murken J, Meitinger T. DHPLC mutation analysis of the hereditary nonpolyposis colon cancer (HNPCC) genes hMLH1 and hMSH2. J Biochem Biophys Methods 2001;47:21-32.
9. Kurzawski G, Safranow K, Suchy J, Chlubek D, Scott RJ, Lubinski J. Mutation analysis of MLH1 and MSH2 genes performed by denaturing high-performance liquid chromatography. J Biochem Biophys Methods 2002;51:89-100.
10. Mihalatos M, Apessos A, Triantafillidis JK, Kosmidis PA, Fountzilas G, Agnantis NJ, Yannoukakos D, Nasioulas G. Evaluation of dHPLC in mutation screening of the APC gene in a Greek FAP cohort. Anticancer Res 2003;23:2691-5.
11. Zhou XL, Eriksson U, Werelius B, Kressner U, Sun XF, Lindblom A. Definition of candidate low risk APC alleles in a Swedish population. Int J Cancer 2004;110:550-7.

12. Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RK, Bisgaard ML, Orntoft TF, Aaltonen LA, Hodgson SV, Thomas HJ, Tomlinson IP. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med 2003;348:791-9.
13. McClay JL, Sugden K, Koch HG, Higuchi S, Craig IW. High-throughput single-nucleotide polymorphism genotyping by fluorescent competitive allele-specific polymerase chain reaction (SNiPTag). Anal Biochem 2002;301:200-6.
14. Sellner LN, Taylor GR. MLPA and MAPH: new techniques for detection of gene deletions. Hum Mutat 2004;23:413-9.
15. Charbonnier F, Olschwang S, Wang Q, Boisson C, Martin C, Buisine MP, Puisieux A, Frebourg T. MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. Cancer Res 2002;62:848-53.
16. Gille JJ, Hogervorst FB, Pals G, Wijnen JT, van Schooten RJ, Dommering CJ, Meijer GA, Craanen ME, Nederlof PM, de Jong D, McElgunn CJ, Schouten JP, Menko FH. Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. Br J Cancer 2002;87:892-7.
17. Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. Hum Mutat 2003;22:258.
18. Meuller J, Kanter-Smoler G, Nygren AO, Errami A, Gronberg H, Holmberg E, Bjork J, Wahlstrom J, Nordling M. Identification of genomic deletions of the APC gene in familial adenomatous polyposis by two independent quantitative techniques. Genet Test 2004;8:248-56.
19. Akrami SM, Dunlop MG, Farrington SM, Frayling IM, Macdonald F, Harvey JF, Armour JA. Screening for exonic copy number mutations at MSH2 and MLH1 by MAPH. Fam Cancer 2005;4:145-149.
20. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 2002;30:e57.
21. Taylor CF, Charlton RS, Burn J, Sheridan E, Taylor GR. Genomic deletions in MSH2 or MLH1 are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. Hum Mutat 2003;22:428-33.
22. Wehner M, Mangold E, Sengteller M, Friedrichs N, Aretz S, Friedl W, Propping P, Pagenstecher C. Hereditary nonpolyposis colorectal cancer: pitfalls in deletion screening in MSH2 and MLH1 genes. Eur J Hum Genet 2005;13:983-6.
23. Ainsworth PJ, Koscinski D, Fraser BP, Stuart JA. Family cancer histories predictive of a high risk of hereditary non-polyposis colorectal cancer associate significantly with a genomic rearrangement in hMSH2 or hMLH1. Clin Genet 2004;66:183-8.
24. Michils G, Tejpar S, Thoelen R, van Cutsem E, Vermeesch JR, Fryns JP, Legius E, Matthijs G. Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): a Belgian study. Hum Mutat 2005;25:125-34.
25. Bunyan DJ, Eccles DM, Sillibourne J, Wilkins E, Thomas NS, Shea-Simonds J, Duncan PJ, Curtis CE, Robinson DO, Harvey JF, Cross NC. Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. Br J Cancer 2004;91:1155-9.
26. Viel A, Petronzelli F, Della Puppa L, Lucci-Cordisco E, Fornasarig M, Pucciarelli S, Rovella V, Quaia M, Ponz de Leon M, Boiocchi M, Genuardi M. Different molecular mechanisms underlie genomic deletions in the MLH1 Gene. Hum Mutat 2002;20:368-74.
27. van der Klift H, Wijnen J, Wagner A, Verkuilen P, Tops C, Otway R, Kohonen-Corish M, Vasen H, Oliani C, Barana D, Moller P, Delozier-Blanchet C, Hutter P, Foulkes W, Lynch H, Burn J, Moslein G, Fodde R. Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC). Genes Chromosomes Cancer 2005;44:123-38.
28. Baudhuin LM, Mai M, French AJ, Kruckeberg KE, Swanson RL, Winters JL, Courteau LK, Thibodeau SN. Analysis of hMLH1 and hMSH2 gene dosage alterations in hereditary nonpolyposis colorectal cancer patients by novel methods. J Mol Diagn 2005;7:226-35.

Summary

Colorectal cancer is the second leading cause of cancer death in the western countries after lung cancer. Colorectal cancer manifests itself after an accumulation of several genetic alterations. These mutations (changes within individual genes) can be either somatic (acquired at a time after conception) or inherited. Inherited mutations are present in all cells of the body, whereas somatic mutations are present only in the originating cell and its descendents. Mutations are triggered by environmental factors such as radiation, chemicals, and viruses, and they occur naturally with increasing age. A DNA repair system operates to keep genes intact. If the repair system is defective, mutations may persist and lead to cancer.

Colorectal cancer consists of a solid malignancy localized in the large intestine and traditionally divided into sporadic and hereditary cases. Sporadic cancer occurs in individuals with no evidence of having hereditary genetic alterations. Somatic mutations tend to cause one or a small number of colon polyps characterized by an abnormal growth of the epithelial cells in the lumen of the intestine. Colonoscopy can detect these polyps and removing them is an effective measure to prevent the progression from polyp to colon cancer. Typically symptoms develop late in the course of the disease, with most sporadic colorectal cancer cases diagnosed after the sixth decade of life. Hereditary colon cancer syndromes are caused by specific inherited mutations. Tumors develop after a somatic genetic alteration in the epithelium of the large intestine inactivates the gene with the inherited mutation (Knudson's two-hit model of carcinogenesis). These syndromes are typically characterized by numerous (sometimes thousands) of colon polyps, although some types of hereditary colon cancers may develop directly from cells of the colon lining without a detectable development of polyps.

The most common forms of heritable colorectal cancer are hereditary non-polyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP), which together account for approximately 5% of all colorectal cancer cases and usually affect multiple members of a family. FAP is caused by an inherited functional mutation of one of the alleles of the adenomatous polyposis coli (APC) gene on chromosome 5. Inactivation of the *APC* gene results in a visible phenotype during early adulthood with the development of hundreds of colorectal polyps. Unless the condition is detected and treated early (which involves removal of the colon) subjects with FAP will develop colon cancer. In addition, these subjects are also at risk of developing other types of cancers, such as those of the thyroid gland, stomach and the duodenum. HNPCC is an inherited predisposition to colorectal cancer occurring predominantly in the right side of the large intestine. Individuals predisposed to HNPCC have a lifetime risk of 80% of developing colorectal cancer, which typically arises after the fourth decade of life (at an earlier age than sporadic colorectal

cancer). Most of these individuals remain asymptomatic until the development of cancer. Genetic tests and colonoscopy may provide an early diagnosis in these subjects, which can lead to prevention or a good treatment. Affected subjects also have an increased risk for cancers of the stomach, small intestine, liver, brain, skin, endometrium, ovary and prostate**.** HNPCC is associated with germline genetic alterations in the mismatch repair (MMR) genes. The primary function of the MMR system is to eliminate single base substitutions and insertion-deletion errors that may arise during DNA replication. The system involves several proteins encoded by a number of genes including *MLH1*, *MLH3*, *MSH2*, *MSH3*, *MSH6*, *PMS1*, and *PMS2*, although only a few of them are implicated in HNPCC. HNPCC is also associated with microsatellite instability in tumor DNA. Microsatellites are short nucleotide repeat sequences present in the coding and/or non-coding sequences of various genes. Alteration in the number of repeats (called microsatellite instability) in the coding region may inactivate the gene. Microsatellite instability is unusual because the MMR system recognizes and repairs inaccurate insertions or deletions of nucleotide repeats during DNA replication ('slippage'). When the MMR system is defective the slippage remains unrepaired in many genes just before cell division. In oncogenes and tumor suppressor genes (which control cell growth) microsatellite instability is presumed to lead to cancer. Currently, microsatellite instability analysis is used as a method to identify HNPCC cases with possible mutations in the MMR genes.

The aim of this thesis was to identify the molecular mechanisms that are involved in the genetic predisposition to colorectal carcinogenesis in HNPCC. Hereby I focused on various mechanisms involved in the constitutional inactivation of MMR genes, such as mutations, methylation of the promoter regions and genetic alterations in the regulatory areas.

The first chapter of this thesis presents an overview of the molecular and clinical aspects of colorectal cancer (Chapter 1.1), including the inherited syndromes HNPCC (Chapter 1.2) and FAP (Chapter 1.3). A brief description of the genetic techniques used in this thesis is reported in the Appendix.

In Chapter 2 linkage analysis was used as a method of identifying individuals at risk of developing FAP but for which no *APC* germline mutations had been detected. In this study 217 individuals from 20 FAP families were analyzed using seven polymorphic markers. The analysis identified the haplotype associated with the disease in 18 families and its segregation in all relatives. An increased risk of developing FAP was established in 10 of the 53 presymptomatic individuals. Further clinical screening tests were suggested to these individuals.

In Chapter 3 a microsatellite instability analysis was used to identify colorectal cancer cases with suspected genetic alterations in the MMR genes. International guidelines classify microsatellite instability phenotype into two distinct groups of tumors depending on the fraction of mutated microsatellite sequences: MSI-L (1 out of 5 selected microsatellites is mutated) and MSI-H (≥2 out of 5). Microsatellite instability analysis on 45 colorectal cancer patients, of which 13 had HNPCC, revealed MSI-H status in 85% of HNPCC patients and 16% of non-HNPCC patients. Thus, microsatellite instability was found to be associated with HNPCC. HNPCC cases with MSI-H were also found to be correlated with the presence of genetic alterations in MMR genes (Chapter 4). A mutational analysis on 16 HNPCC families, of which 46% of the individuals had MSI-H tumors, identified eight genetic alterations in *MLH1* and *MSH2*.

Chapter 5 reports a first case of promoter methylation as an epigenetic alteration at a constitutional level. Methylation analysis of the *MLH1* promoter region was performed on 14 HNPCC cases with MSI-H but without germline mutations in the MMR genes, using a methylation-specific polymerase (MSP) chain reaction assay. One case was identified to have one methylated and one unmethylated *MLH1* allele in blood DNA, and only the methylated allele in tumor tissue. Analysis of the single nucleotide polymorphisms (SNPs) in the promoter region showed that the blood DNA was heterozygous for the polymorphism (G/A) whereas the tumor DNA was homozygous (G), indicating a loss of heterozygosity (LOH) resulting in a loss of the unmethylated allele at the somatic level. Inactivation of both *MLH1* alleles resulted in an absence of MLH1 protein as confirmed by immunohistochemistry. In the last few years other investigators have confirmed these findings. Thus, constitutional hemi-methylation is a new mechanism for monoallelic inactivation of the *MLH1* gene in HNPCC and sporadic early-onset colon cancer patients with MSI-H.

Chapter 6 describes a functional analysis of the 5'-flanking region of the *MSH6* gene and its regulation by seven Sp1 transcription factor binding sites. Identification of two SNPs in the *MSH6* promoter region in cancer patients showed that two of the seven Sp1 transcription factor binding sites were inactivated, resulting in a reduced promoter activity. In addition, *in vitro* experiments revealed that an *MSH6* promoter containing two SNPs was more sensitive to inactivation by DNA methylation. Two different polymorphic variants were identified in 16% of Caucasians and 3% of a diversity population, suggesting that the human population contains individuals with different levels of *MSH6* promoter activity. We also showed that mutations in the regulatory promoter region can reduce the binding of transcription factors such as Sp1, resulting in a reduced promoter activity and possible de novo methylation. Thus, we provide a new mechanism that can cause partial

inactivation of an MMR gene and reduced constitutional gene expression rather than alterations in the structure of the protein.

In conclusion, in this thesis three different mechanisms are identified that can lead to the constitutional inactivation of MMR genes in HNPCC, including a high incidence of point mutations. Most interestingly, two novel mechanisms were identified that occurred at a low incidence: (i) methylation of the *MLH1* gene, and (ii) SNPs in the transcription factor sites of the *MSH6* gene.

Nederlandse samenvatting

Dikkedarmkanker is na longkanker de meest voorkomende doodsoorzaak ten gevolge van kanker in de westerse wereld. Dikkedarmkanker manifesteert zich na een accumulatie van verscheidene genetische veranderingen. Mutaties (veranderingen in individuele genen) kunnen somatisch (opgedaan na de conceptie) of geërfd zijn. Geërfde mutaties bevinden zich in alle cellen van het lichaam, somatische mutaties bevinden zich daarentegen alleen in de cel waarin de mutatie plaastvindt en de afstammelingen daarvan. Mutaties kunnen worden teweeggebracht door omgevingsfactoren zoals straling, chemicaliën en virussen, en doen zich van nature voor bij het ouder worden. Een DNA reparatiesysteem zorgt ervoor dat genen intact blijven. Als dit systeem defect is kunnen mutaties aanwezig blijven en tot kanker leiden.

Dikkedarmkanker bestaat uit een solide tumor in de dikke darm en wordt traditioneel onderverdeeld in sporadische en erfelijke gevallen. Bij individuen bij wie geen aanwijzingen van erfelijke genetische veranderingen zijn aangetroffen spreekt men van sporadische kanker. Somatische mutaties veroorzaken gewoonlijk een of een klein aantal poliepen in de karteldarm (colon) die gekenmerkt worden door een abnormale groei van de epitheelcellen in het lumen van de darm. Coloscopie kan deze poliepen ontdekken. Het verwijderen van de poliepen is een effectieve maatregel die voorkomt dat deze verder uitgroeien tot dikkedarmkanker. De symptomen ontwikkelen zich gewoonlijk laat in het verloop van de ziekte, waarbij de meeste sporadische dikkedarmkanker gevallen worden vastgesteld na het 60^{e} levensjaar. Erfelijke dikkedarmkanker syndromen worden veroorzaakt door specifieke geërfde mutaties. Tumoren ontstaan nadat een somatische genetische verandering in het epitheel van de dikke darm het gen met de geërfde mutatie inactiveert (Knudson's two-hit carcinogenesis model). Deze syndromen worden gewoonlijk gekenmerkt door vele (soms duizenden) poliepen in de dikke darm, alhoewel sommige soorten erfelijke dikkedarmkanker zich direct kunnen ontwikkelen uit de epitheelcellen van de dikke darm zonder dat een poliepontwikkeling detecteerbaar is.

De meest voorkomende vormen van erfelijke dikkedarmkanker zijn hereditair non-polyposis colorectaal carcinoom (HNPCC) en familiaire adenomateuze polyposis (FAP), welke samen ongeveer 5% van alle dikkedarmkanker gevallen voor rekening nemen en gebruikelijk meerdere familieleden raken. FAP wordt veroorzaakt door een geërfde functionele mutatie van een van de kopieën van het adenomatous polyposis coli (APC) gen op chromosoom 5. Inactivatie van het *APC* gen resulteert in een zichtbaar fenotype tijdens de vroege volwassenheid dat gekenmerkt wordt door een ontwikkeling van honderden poliepen in de dikke darm. Tenzij de conditie vroeg ontdekt en behandeld wordt (gepaard gaande met het verwijderen van de dikke darm) zullen personen met FAP dikkedarmkanker ontwikkelen. Deze personen lopen tevens het risico om andere soorten

kanker te ontwikkelen, onder andere van de schildklier, maag en de twaalfvingerige darm. HNPCC is een erfelijke vorm van dikkedarmkanker die voornamelijk aangetroffen wordt in de rechter zijde van de dikke darm. Individuen met een aanleg voor HNPCC lopen een levenstijd risico van 80% om dikkedarmkanker te ontwikkelen, welke zich gewoonlijk voordoet na het 40^{e} levensjaar (op een vroegere leeftijd dan sporadische dikkedarmkanker). De meesten van deze personen blijven asymptomatisch tot aan de ontwikkeling van kanker. Door genetische testen en coloscopie kan een vroege diagnose gesteld worden die kan leiden tot preventie of een goede behandeling. Individuen met een aanleg voor HNPCC lopen ook een verhoogd risico op kanker van de maag, dunne darm, lever, hersenen, huid, endometrium, eierstok en prostaat. HNPCC is geassocieerd met kiemlijn genetische veranderingen in de mismatch repair (MMR) genen. De primaire functie van het MMR systeem is het elimineren van base substituties en insertie-deletie fouten die ontstaan tijdens DNA replicatie. Het systeem omvat verscheidene eiwitten die gecodeerd worden door een aantal genen, waaronder *MLH1*, *MLH3*, *MSH2*, *MSH3*, *MSH6*, *PMS1* en *PMS2*, waarvan een aantal geimpliceerd zijn in HNPCC. HNPCC is ook geassocieerd met microsatelliet instabiliteit in tumor DNA. Microsatellieten zijn korte sequenties van repeterende nucleotiden die aanwezig zijn in de coderende en/of niet-coderende sequenties van verscheidene genen. Een verandering in het aantal repetities ('microsatelliet instabiliteit') in het coderende gebied kan het gen inactiveren. Microsatelliet instabiliteit is ongebruikelijk omdat het MMR systeem onnauwkeurige inserties of deleties van nucleotide repetities tijdens DNA replicatie ('slippage') herkent en herstelt. Als het MMR systeem defect is worden deze fouten echter niet hersteld, wat in oncogenen en tumor-suppressie genen (die de celgroei controlleren) tot kanker kan leiden. Microsatelliet instabiliteit analyse wordt daarom tegenwoordig gebruikt als een methode om HNPCC gevallen te identificeren met mogelijke mutaties in de MMR genen.

Het doel van dit proefschrift was om de moleculaire mechanismen die betrokken zijn bij de genetische aanleg voor colorectale carcinogenesis in HNPCC te identificeren. Hierbij lag de nadruk op de verscheidene mechanismen die betrokken zijn bij de constitutionele inactivatie van MMR genen, zoals mutaties, methylatie van de promotor gebieden en genetische veranderingen in de regulatie gebieden.

In het eerste hoofdstuk van dit proefschrift wordt een overzicht gepresenteerd van de moleculaire en klinische aspecten van dikkedarmkanker (Hoofdstuk 1.1), inclusief de erfelijke syndromen HNPCC (Hoofdstuk 1.2) en FAP (Hoofdstuk 1.3). De genetische technieken die gebruikt worden in dit proefschrift worden kort beschreven in de Appendix.

In Hoofdstuk 2 wordt linkage analyse gebruikt als een methode om personen te identificeren die risico lopen om FAP te ontwikkelen, maar bij wie geen *APC* kiemlijn mutaties zijn gevonden. In deze studie werden 217 personen uit 20 FAP families geanalyseerd met behulp van zeven polymorfe markers. De analyse identificeerde het haplotype dat geassocieerd is met de ziekte in 18 families en de segregatie daarvan in alle familieleden. Een verhoogd risico om FAP te ontwikkelen werd gevonden in 10 van de 53 presymptomatische individuen. Verdere klinische screening testen werden aanbevolen aan deze personen.

In Hoofdstuk 3 wordt een microsatelliet instabiliteit analyse gebruikt om dikkedarmkanker gevallen met vermoedelijke genetische veranderingen in de MMR genen te identificeren. Het microsatelliet instabiliteit fenotype wordt volgens internationale richtlijnen in twee verschillende groepen tumoren onderverdeeld afhankelijk van de fractie van gemuteerde microsatelliet sequenties: MSI-L (1 uit 5 geselecteerde microsatellieten is gemuteerd) en MSI-H (≥2 uit 5). Microsatelliet instabiliteit analyse van 45 dikkedarmkanker patienten, waarvan 13 met HNPCC, toonde MSI-H status aan in 85% van de HNPCC patienten en 16% van de niet-HNPCC patienten. Microsatelliet instabiliteit was dus geassocieerd met HNPCC. HNPCC gevallen met MSI-H waren ook gecorreleerd met de aanwezigheid van genetische veranderingen in de MMR genen (Hoofdstuk 4). Een mutatie analyse van 16 HNPCC families, waarin 46% van de individuen MSI-H tumoren hadden, identificeerde acht genetische veranderingen in *MLH1* en *MSH2*.

Hoofdstuk 5 rapporteert een eerste geval van promotor methylatie als een epigenetische verandering op een constitutioneel niveau. Methylatie analyse van het *MLH1* promotorgebied werd uitgevoerd op 14 HNPCC gevallen met MSI-H maar zonder kiemlijn mutaties in de MMR genen, met behulp van een methylatie-specifieke polymerase (MSP) kettingreactie assay. Een geval was geidentificeerd met een gemethyleerd en een ongemethyleerd *MLH1* allel in DNA uit het bloed, en alleen het gemethyleerde allel in tumor weefsel. Analyse van de single nucleotide polymorfismen (SNPs) in het promotorgebied toonde aan dat het bloed DNA heterozygoot was voor het polymorfisme (G/A) terwijl het tumor DNA homozygoot (G) was. Dit duidt op een loss of heterozygosity (LOH) die resulteert in een verlies van het ongemethyleerde allel op het somatische niveau. Immunohistochemie bevestigde dat inactivatie van beide *MLH1* allelen resulteerde in de afwezigheid van MLH1 eiwit. Andere onderzoekers hebben de afgelopen jaren deze bevindingen bevestigd. Constitutionele hemi-methylatie is dus een nieuw mechanisme voor monoallelische inactivatie van het *MLH1* gen in HNPCC en sporadische early-onset dikkedarmkanker patienten met MSI-H.

Hoofdstuk 6 beschrijft een functionele analyse van het 5'-flankerend gebied van het *MSH6* gen en de regulatie daarvan door zeven Sp1 transcriptiefactor bindingsplaatsen. De identificatie van twee SNPs in het *MSH6* promotor gebied in kanker patienten toonde aan dat twee van de zeven Sp1 transcriptiefactor bindingsplaatsen geinactiveerd waren, wat resulteerde in een verlaagde promotor activiteit. In vitro experimenten toonden bovendien aan dat een *MSH6* promotor met twee SNPs sensitiever is voor inactivatie door DNA methylatie. Twee verschillende polymorfe varianten werden geidentificeerd in 16% van een Kaukasische populatie en 3% van een gemengde populatie, wat suggereert dat de menselijke populatie individuen bevat met verschillende niveaus van *MSH6* promotor activiteit. We hebben ook aangetoond dat mutaties in het regulerende promotorgebied het binden van transcriptiefactoren zoals Sp1 kan verminderen, wat resulteert in een verlaagde promotor activiteit en mogelijke de novo methylatie. We verschaffen dus een nieuw mechanisme dat een gedeeltelijke inactivatie van een MMR gen kan veroorzaken en een verlaagde constitutionele gen expressie in plaats van veranderingen in de structuur van het eiwit.

Tot besluit, in dit proefschrift zijn drie verschillende mechanismen geidentificeerd die kunnen leiden tot de constitutionele inactivatie van MMR genen in HNPCC. Een veelvoorkomend mechanismse betreft verscheidene punt mutaties. Twee nieuwe maar weinig voorkomende mechanismen zijn (i) methylatie van het *MLH1* gen en (ii) SNPs in de transcriptiefactor bindingsplaatsen van het *MSH6* gen.

Acknowledgements

I have worked in different laboratories where I met wonderful people. Each of them left something in my heart and increased my scientific knowledge by countless discussions. I would like to express my gratitude to all colleagues, friends and family who contributed to the work described in this thesis or just supported me during these years. Without their help I would not have reached this moment.

First, I would like to thank my promotor, Professor Hans Clevers, who supported and helped me realize this important step of my life. I wish to express my special thanks to Professor Riccardo Fodde, who has been a wonderful mentor during the last two years. I am also thankful to the members of the committee for their time and interest in my work. I especially thank Janny van Eldik, who has been very kind and helped me from the first moment we met.

A big part of this thesis was performed at the Ludwig Institute for Cancer Research at the University of California, San Diego. I am sincerely grateful to my supervisor, Professor Richard Kolodner, for giving me the opportunity to work with him and his help and patience during the years in his lab. I would also like to thank my colleagues and friends in the lab: Cecilia, Chris, Dan, Ellen, Hal, James, Jill, Marc, Martin, Sandra, Scarlet, Stephanie and Vincent. Thanks for all these good sunny moments of scientific or personal discussions in the lab or during the coffee break at Mandeville café. I could also not forget the kind help of Ann Montgomery, who was always there when I needed it.

The first part of my study began in Italy at the Molecular Pathology Institute at the Hospital Policlinico in Milan. I thank Dr. Giuseppe Colucci for introducing me to the field of cancer genetics and providing me with enthusiasm and interest in this field. I also would like to thank Professor Giuseppe Simoni and my other colleagues. In the last few years of my studies in Italy, I worked at the National Cancer Institute in Milan. Here I would like to thank Dr. Paolo Radice for his support and friendship during the years that we worked together. I would like to thank Dr. Marco Pierotti, Dr. Lucio Bertario and Paola Sala for the scientific discussions and advice. Thank you to all my colleagues and especially Patrizia Mondini for the long-time friendship across the ocean. In addition, I want to thank Prof. Nadia Ranzani and Dr. Daniele Calistri for our collaboration.

I would like to thank all my colleagues at Genmab in Utrecht, especially Adinda, Adrie, Annemiek, Aran, Erik, Ibraham, Karin, Luus, Marcel, Marleen, Marouska, Michel, Mina, Patrick, Tom, Touria and Wendy. Thanks for sharing with me great moments full of talk, laughs and scientific discussions. My personal gratitude goes to Jeroen Lammerts van Bueren for his great advice, and Professor Jan van de Winkel for his support and enthusiasm for my thesis research.

I would like to thank my many colleagues at DSM Anti-Infectives in Delft for a great time and their friendship. Special thanks to Richard, Denise, Daan, Paul, Ben and Rob for their encouragement during the last bits and pieces of writing this thesis.

Big thanks to all my friends around the world for sharing all these happy and wonderful moments with me during my years abroad, especially Ana, Bea, Johan, Tomas, Anya, Luca, Jasmina, Laura, Giusy, Barbara, Nancy, Geert, Martine, Bridget, Henri, Ron, Paola, Marcin, Valeria and Martin.

A special thanks to my paranimf, Johan de Rooij. You stood by me when I needed it. Thanks Johan, you have been a wonderful friend.

I would like to thank my mother, my father and my sister Donatella for always being with me and offering a hand during this long journey. My family in law, Dini, Aad, Guido and Frank, thank you for teaching me to be part of a Dutch family. Thank you Luca, Mattia and Rina for supporting me and loving who I love.

Lastly, I would like to thank the most important person, my husband Arthur. Thank you for your love and your unlimited patience. Thank you for being always with me, drying my tears, making me laugh and strongly supporting me when I needed it. Thank you for being what “I am not” and making me feel special. I would have never made it without you.

Now, we are ready for a new life in Berlin.

List of publications

Gazzoli I, De Andreis C, Sirchia SM, Sala P, Rossetti C, Bertario L, Colucci G. Molecular screening of families affected by familial adenomatous polyposis (FAP). *J Med Screen. 1996; 3(4):195-9.*

Sirchia SM, Garagiola I, De Andreis C, Gazzoli I, Gramegna M, Colucci G. Characterization of four microsatellites in an Italian population and their application to paternity testing. *Mol Cell Probes. 1996 Apr; 10(2):155-8.*

Pensotti V, Radice P, Presciuttini S, Calistri D, Gazzoli I, Grimalt Perez A, Mondini P, Buonsanti G, Sala P, Rossetti C, Ranzani GN, Bertario L, Pierotti MA. Mean age of tumor onset in hereditary nonpolyposis colorectal cancer (HNPCC) families correlates with the presence of mutations in DNA mismatch repair genes. *Genes Chromosomes Cancer. 1997 Jul; 19(3):135-42*.

Calistri D, Presciuttini S, Buonsanti G, Radice P, Gazzoli I, Pensotti V, Sala P, Eboli M, Andreola S, Russo A, Pierotti M, Bertario L, Ranzani GN. Microsatellite instability in colorectal-cancer patients with suspected genetic predisposition. *Int J Cancer. 2000 Jan 20; 89(1):87-91*.

Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res. 2002 Jul 15; 62(14):3925-8.*

Gazzoli I, Kolodner RD. Regulation of the human MSH6 gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms. *Mol Cell Biol. 2003 Nov; 23(22):7992-8007.*

Bertario L, Russo A, Sala P, Alberici P, Bassi C, Gazzoli I, Montefusco C, Torrini M, Mareni C, Fornasarig P, Santarosa M, Viel A, Benatti P, Pedroni M, Ponz de Leon M, Lucci-Cordisco E, Genuardi M, Messerini L, Casale V, De Lellis L, Cama A, Pierotti MA and Radice P. Prognostic Relevance of *MLH1* and *MSH2* Mutations in Hereditary Nonpolyposis Colorectal Cancer Patients Affected with Colorectal Cancer. *Submitted for publication.*

Curriculum vitae

The author of this thesis was born on March 5, 1970 in Milan, Italy. After finishing high school in 1988 she studied biological sciences at the University of Pavia, Italy, with a specialization in molecular biology. In 1993 she received her Master's degree in biological sciences (dottore in scienze biologiche). The same year she started to work as a research associate at the Molecular Pathology Institute at the Ospedale Maggiore Policlinico in Milan, where she used molecular techniques to investigate families affected with familial adenomatous polyposis (FAP), under the supervision of dr. G. Simoni and dr. G. Colucci. Subsequently from 1995 to 1999 she worked at the experimental oncology division of the National Cancer Institute (INT) in Milan, under the guidance of dr. M. Pierotti and dr. P. Radice. During this period she investigated the genetic alterations of mismatch repair (MMR) genes in hereditary nonpolyposis colorectal cancer (HNPCC). Results of these studies are presented in this thesis. In 1999 she moved to the USA and continued her research on colorectal cancer in the laboratory of dr. R. Kolodner at the Ludwig Institute for Cancer Research at the University of California, San Diego. The focus of her research here was on identifying the inactivation mechanisms of MMR genes operating at a constitutional level in HNPCC. Results of these studies are reported in this thesis. After 5 years, at the end of 2004, she moved to the Netherlands where she worked for 9 months as a research associate at Genmab in Utrecht. Currently, she works as a research associate at DSM Anti-Infectives in Delft, the Netherlands.

List of abbreviations

A	Adenine
(A)n	Adenosine mononucleotide repeats
ACI	Amsterdam criteria I
ACII	Amsterdam criteria II
APC	Adenomatous polyposis coli gene
ATP	Adenosine triphosphate
BAX	Bcl2-associated X protein
BRCA1, 2	Breast cancer gene
CpG	Cytidine and guanosine dinucleotide
(CA)n	Cytidine and adenosine mononucleotides repeats
CCD	Charged coupled device
CHRPE	Congenital hypertrophies of the retinal epithelium
CNS	Central nervous system
CRC	Colorectal cancer
DGGE	Denaturing gradient gel electrophoresis
DHPLC	Denaturing high-performance liquid chromatograph
DNA	Deoxyribonucleic Acid
ddNTPs	Dideoxynucleotides
E2F4	E2F transcription factor 4 gene
EGFR	Epidermal growth factor receptor gene
FAP	Familial adenomatous polyposis
gDNA	genomic DNA
(GT)n	Guanosine and thymidine mononucleotides repeats
HNPCC	Hereditary non-polyposis colorectal cancer
ICG-HNPCC	International Collaborative Group-Hereditary Nonpolyposis Colorectal Cancer
IDLs	Insertion/deletion loops
IHC	Immunohistochemistry
IVSP	*In vitro* synthesized protein assay
LOH	Loss of heterozygosity
MAP	mutY human homologue (MYH) associated polyposis

MAPH	Multiplex amplification and probe hybridization
MCR	Mutation cluster region
hMLH 1, 3	MutL homolog 1, 3
MLPA	Multiplex ligation-dependent probe amplification
mRNA	messager Ribonucleic acid
MMP	Microsatellite mutator phenotype
MMR	Mismatch repair
MMS	Microsatellite stable
hMSH 2, 3, 6	MutS homolog 2, 3, 6
MSI	Microsatellite instability
MSI-H	Microsatellite instability High
MSI-L	Microsatellite instability Low
hMutL α	Heterodimer of hMLH1 and hPMS2
hMutL β	Heterodimer of hMLH1 and hMLH3 (or hPMS1)
hMutSα	Heterodimer of hMSH2 and hMSH6
hMutSβ	Heterodimer of hMSH2 and hMSH3
NCI	National Cancer Institute
PCR	Polymerase chain reaction
hPMS1	Postmeiotic segregation increased 1, homolog yMLH2
hPMS2	Postmeiotic segregation increased 2, homolog yPMS1
PTEN	Phosphatase and tensin homolog
PTT	Protein truncation test
Rb	Retinoblastoma gene
RER+	Replication error positive
RNA	Ribonucleic acid
SSCP	Single-strand conformation polymorphisms
SNiPTag	Single-Nucleotide Polymorphism Genotyping by Fluorescent Competitive Allele-Specific Polymerase Chain Reaction
(T)n	Thymidine mononucleotide repeats
TATA	Thymidine and adenosine mononucleotide repeats
TGF-β RII	Transforming growth factor-β type II receptor