DNA mismatch repair defects: role in colorectal carcinogenesis

Sandrine Jacob, Françoise Praz *

CNRS, Unité Propre de Recherches 2169, 'Genetic Instability and Cancer', 7, rue Guy-Môquet, 94800 Villejuif, France

Received 30 November 2001; accepted 30 November 2001

Abstract

The inactivation of the DNA mismatch repair (MMR) system, which is associated with the predisposition to the hereditary non-polyposis colorectal cancer (HNPCC), has also been documented in nearly 20% of the sporadic colorectal cancers. These tumors are characterized by a high frequency of microsatellite instability (MSI+ phenotype), resulting from the accumulation of small insertions or deletions that frequently arise during replication of these short repeated sequences. A germline mutation of one of the two major MMR genes (hMSH2 or hMLH1) is found in half to two-thirds of the patients with HNPCC, whereas in sporadic cases hypermethylation of the hMLH1 promoter is the major cause of the MMR defect. Germline mutations in hMSH6 are rare and rather confer predisposition to late-onset familial colorectal cancer, and frequent extracolonic tumors. Yet, the genetic background of a number of HNPCC patients remains unexplained, indicating that other genes participate in MMR and play important roles in cancer susceptibility. The tumor-suppressor genes that are potential targets for the MSI-driven mutations because they contain hypermutable repeated sequences are likely to contribute to the etiology and tissue specificity of the MSI-associated carcinogenesis. Because the prognoisis and the chemosensitivity of the MSI+ colorectal tumors differ from those without instability, the determination of the MSI phenotype is expected to improve the clinical management of patients. This review gives an overview of various aspects of the biochemistry and genetics of the DNA mismatch repair system, with particular emphasis in its role in colorectal carcinogenesis. © 2002 Société française de biochimie et biologie moléculaire / Éditions scientifiques et médicales Elsevier SAS

Keywords: DNA mismatch repair; Colorectal carcinogenesis; Tumor-suppressor gene; Methylation

1. Introduction

The mismatch repair system (MMR) was originally identified in bacteria, in which its inactivation results in an increase in the rate of spontaneous mutations owing to the inability to repair replication errors [1–5]. The importance of this DNA repair system has clearly emerged with the demonstration in the early 1990s that inactivation of the corresponding human pathway is the primary cause of hereditary nonpolyposis colorectal cancers (HNPCC) (reviewed in [6–10]). The MMR system is best known for its role in the post-replicative repair of the errors made by DNA polymerases that have escaped proofreading. In addition to correcting misincorporated nucleotides, MMR proteins have also been implicated in double-strand break repair and recombination [11]. In particular, MMR recognizes the mismatches in heteroduplex recombination intermediates preventing completion of recombination between diverged sequences, thus promoting genetic stability [12,13]. MMR also participates in an additional DNA repair process, the transcription-coupled repair pathway [14,15]. Last but not least, some MMR components have been involved in cell-cycle regulation and the p53-dependent apoptotic response to a variety of DNA damage [16–22]. These last observations may be particularly relevant to the chemotherapeutic treatment of colorectal cancers deficient in mismatch repair [18,23–25].

2. DNA mismatch repair in Escherichia coli

The E. coli DNA mismatch repair system, also known as long patch mismatch repair, has been completely reconstr...
tuted in vitro (Fig. 1). It is mediated by three specialized proteins: MutS, MutL, and MutH [26–28]. Only a short overview of the bacterial MMR is given here because reviews dealing with the genetic and biochemical aspects of MMR in E. coli have been recently published (reviewed in [11,29,30]). The first step in the correction of replication errors by MMR consists in efficient recognition of the base/base mismatches and small insertion or deletion loops (IDL), and is performed by MutS. MutS is an ATPase acting as a homodimer that translocates along DNA and promotes DNA loop formation [31] (reviewed in [4,5,11]). The recent high-resolution crystal structures of MutS of E. coli and its Thermus aquaticus homolog led to the description of how MutS recognizes the various replication errors, and revealed that the two monomers of MutS are asymmetric, not only in the ATPase domains but also in the DNA-binding domains [32,33]. After loading onto the DNA, MutS undergoes a conformational change, which, in the presence of ATP, permits its interaction with MutL [34]. MutL, another ATPase acting as a homodimer, couples mismatch recognition to further steps resulting in the removal of a large newly replicated DNA fragment containing the mismatch [35,36]. The discrimination between the template and the newly synthesized strands is determined by adenine methylation at GATC sequences by the dam methylase, a process that occurs shortly after replication [26]. MutH is an endonuclease that can nick the nascent strand at hemimethylated GATC sequences [37]. MutL stimulates the endonuclease activity of MutH when engaged in a complex with MutS [38]. Furthermore, MutL enhances the ATP hydrolysis-dependent translocation of MutS along the DNA in search of the closest unmethylated GATC site bound by MutH. Excision of the fragment containing the mismatch requires the MutU/UvrD (DNA helicase II) gene product helicase, which is loaded onto the nick by MutL [39–41]. Because MutH can nick DNA on either side of a mismatch, excision requires exonucleases that have either 3′ to 5′ or 5′ to 3′ directionality, depending on whether the MutH excision occurs 3′ or 5′ to the mispair [27,42–44]. At least four exonucleases, recJ, ExoVII, ExoI, and ExoX have been shown to participate in MMR [44,45]. A large fragment of up to several hundred nucleotides is removed, resynthesized by the replicating DNA polymerase III holoenzyme and ligated [28,46].

### 3. Eukaryotic DNA mismatch repair systems

The key components of MMR are highly conserved from bacteria to mammals. This evolutionary conservation has been of great help for defining the eukaryotic MMR systems, in which homologs of all bacterial genes have been identified with the exception of the endonuclease MutH. A great deal of our knowledge of the human system is based on the observations made in Saccharomyces cerevisiae.

#### 3.1. The MutS-related complexes

Yeast has six homologs of MutS (MSH1–MSH6), whereas mammals lack MSH1 homologs. Four homologs of MutL (MLH1–MLH3 and PMS1) have been identified in both yeast and mammals. These multiple homologs of MutS and MutL play specialized or partly redundant functions in yeast and mammals (Tables 1 and 2) [47–58]. With the exception of MSH4 and MSH5, which promote crossing-over and chromosome synopsis during meiosis, all MSH homologs participate in mutation avoidance [48,50,52–58]. MSH1 is involved in the repair and maintenance of mitochondrial DNA, whereas MSH2, MSH3, and MSH6 are required for the stability of nuclear DNA [47,59–62].

Genetic and biochemical data show that MSH2 can form heteroduplexes with either MSH6 or MSH3, leading to the formation of MutSα and MutSβ, respectively [62–65]. It is worth noting that in the complex formed between MSH2
and MSH6, only the latter subunit contacts the mismatched substrate, as in the case of bacterial MutS in which the two monomers are asymmetric [32,33,66]. The studies performed in yeast have demonstrated that the MutSα complex is able to bind to most base–base mismatches (with the exception of CC mismatches), and to loops of one or a few nucleotides, whereas MutSβ preferentially repairs heteroduplexes with two or more extrahelical bases [51,62,67,68].

In human cells, the mismatch recognition specificities of MutSα and MutSβ are very similar to those of their yeast counterparts. The mismatch recognition is mainly performed by MutSα, which is present at much higher levels than MutSβ [63,64,67,68]. Because the MSH3 and MSH6 functions are partially redundant, disruption of either gene leads to a weak mutator phenotype, whereas the cells deficient in both MSH3 and MSH6 have a high mutation rate, similar to that observed in the MSH2-deficient cells [62,71]. The inactivation of MSH2 is responsible for a characteristic phenotype referred to as ‘Replication ERror’ (RER), or ‘MicroSatellite Instability’ (MSI) because microsatellite sequences are repetitive sequences of one to a few base pairs that are particularly prone to polymerase slippage and errors.

### 3.2. The MutL-related complexes

The existence of four MutL homologs, MLH1–MLH3 and PMS1 has been reported in yeast (Table 2) [72,73]. In human, in addition to the four MutL homologs, namely hMLH1, hMLH3, hPMS1 and hPMS2, a cluster of hPMS2-like genes have been localized on chromosome 7 [74,75]. The three different heterodimeric hMutL-related complexes all contain hMLH1, and it has recently been established that all three hMLH1 partners, hMLH3, hPMS1 and hPMS2, interact with the same region of hMLH1 [76]. The major component in MMR, MutLα, is composed of MLH1 and hPMS2, a cluster of hPMS2-like genes have been localized on chromosome 7 [74,75]. The three different heterodimeric hMutL-related complexes all contain hMLH1, and it has recently been established that all three hMLH1 partners, hMLH3, hPMS1 and hPMS2, interact with the same region of hMLH1 [76]. The major component in MMR, MutLα, is composed of MLH1 and hPMS1 in yeast, and of hMLH1 and hPMS2, the closest homolog of S. cerevisiae PMS1, in human cells [77–79]. Yeast MutLα, as well as hPMS2, has structural and functional properties of the GHL family of ATPases which lack the conventional ATPase signature motif, the Walker A motif [80,81]. In yeast, ATPase motifs of both MLH1 and

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Eukaryotic MutS homologs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast Mammals</td>
<td>Functions</td>
</tr>
<tr>
<td>MSH1</td>
<td>Mutation avoidance in mitochondria</td>
</tr>
<tr>
<td>MSH2 MSH2</td>
<td>Forms heterodimers with MSH3 and MSH6</td>
</tr>
<tr>
<td></td>
<td>• Induces apoptosis in response to certain DNA damage</td>
</tr>
<tr>
<td></td>
<td>• Participates in transcription-coupled repair</td>
</tr>
<tr>
<td></td>
<td>• Facilitates resolution of Holliday junctions</td>
</tr>
<tr>
<td>MSH3 MSH3</td>
<td>Forms heterodimers with MSH2 = MutSβ</td>
</tr>
<tr>
<td></td>
<td>• Repairs insertion/deletion ≥ 2bp</td>
</tr>
<tr>
<td></td>
<td>• Repairs mismatches in recombination intermediates</td>
</tr>
<tr>
<td></td>
<td>• Removes nonhomologous tails during recombination</td>
</tr>
<tr>
<td></td>
<td>• Inhibits recombination between divergent sequences</td>
</tr>
<tr>
<td>MSH4 MSH4</td>
<td>Forms heterodimers with MSH5</td>
</tr>
<tr>
<td></td>
<td>• No function in postreplicative repair</td>
</tr>
<tr>
<td></td>
<td>• Promotes meiotic crossing-over and is involved in chromosome synopsis</td>
</tr>
<tr>
<td>MSH5 MSH5</td>
<td>Forms heterodimers with MSH</td>
</tr>
<tr>
<td></td>
<td>• No function in postreplicative repair</td>
</tr>
<tr>
<td></td>
<td>• Promotes meiotic crossing-over and is involved in chromosome synopsis</td>
</tr>
<tr>
<td>MSH6 MSH6</td>
<td>Forms heterodimers with MLH1 = MutSα</td>
</tr>
<tr>
<td></td>
<td>• Repairs base-base mismatches</td>
</tr>
<tr>
<td></td>
<td>• Repairs insertion/deletion ≥ 1bp</td>
</tr>
<tr>
<td></td>
<td>• Repairs mismatches in recombination intermediates</td>
</tr>
<tr>
<td></td>
<td>• Inhibits recombination between divergent sequences</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Eukaryotic MutL homologs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast Mammals</td>
<td>Functions</td>
</tr>
<tr>
<td>MLH1/PMS2 MLH1</td>
<td>Forms heterodimers with yPMS1/hPMS2, yMLH2/hPMS1, MLH3</td>
</tr>
<tr>
<td></td>
<td>• Induces apoptosis in response to certain DNA damage</td>
</tr>
<tr>
<td></td>
<td>• Promotes chromosome synopsis in meiosis</td>
</tr>
<tr>
<td>PMS1 PMS2</td>
<td>Forms heterodimers with MLH1 = MutLα</td>
</tr>
<tr>
<td></td>
<td>• Acts in concert with MutSα or MutSβ</td>
</tr>
<tr>
<td></td>
<td>• Promotes chromosome synopsis in meiosis</td>
</tr>
<tr>
<td>MLH2 PMS1</td>
<td>Forms heterodimers with MLH1</td>
</tr>
<tr>
<td>MLH3 MLH3</td>
<td>Forms heterodimers with MLH1 = MutLβ</td>
</tr>
<tr>
<td></td>
<td>• Acts in concert with MutLβ</td>
</tr>
<tr>
<td></td>
<td>• Repairs insertion/deletion ≥ 2bp</td>
</tr>
</tbody>
</table>
PMS1 are required for MMR activity [81]. The ATP binding induces a conformational change of MutLα, which promotes the heterodimerization between the N-terminal regions of MLH1 and PMS1 [81]. Human PMS2 has recently been shown to have both an ATPase and DNA-binding activity in its N-terminal fragment [80]. Human PMS1 is more closely related to yeast MLH2. Although hPMS1 and yMLH2 are able to bind MLH1 with a high affinity, their involvement in post-replicative mismatch repair remains to be demonstrated [79,82]. The recently identified human MLH3 gene is most similar to yeast MLH3 [83]. In yeast, MLH3 is able to form a complex with MLH1, and has been shown to act in concert with MutSβ to repair a proportion of the IDL, thus decreasing the rate of frameshift mutations [84]. Although it was shown that cells stably expressing a dominant-negative MLH3 mutant protein display microsatellite instability, the physiologic contribution of the IDL, resulting from MLH1 deficiency leads to a severe phenotype, similar to that of MSH2-deficient cells. The MSI phenotype resulting from MLH1 deficiency is characterized by a tremendous increase in base–base mismatches, as well as frameshift mutations resulting from unrepairred IDL.

3.3. The discrimination of the newly synthesized strand and the processing steps of MMR

Once bound to mismatches, the MutLα complex is able to interact with numerous factors, consistent with the assembly of a higher-order complex that is involved in the excision of a large fragment of the newly synthesized DNA strand containing the mismatch (reviewed in [8,11,85]). In organisms other than E. coli, the signal that allows to discriminate the newly synthesized strand from the template is still doubtful, but does not involve DNA methylation [86]. The proliferating cell nuclear antigen (PCNA) which is essential for DNA replication, where it acts as a processivity factor has also been implicated in MMR before and during the DNA synthesis step and could be involved in this process. Both MutSα and MutSβ are able to bind to PCNA via N-terminal sequence motifs that are present in MSH6 and MSH3 subunits, but not in MSH2 [61,87,88]. Colocalization of PCNA and MSH6 or MSH3 to replication foci implies an intimate connection between replication and mismatch correction [61]. Although the biochemical basis for the role of PCNA early in MMR is not fully understood yet, PCNA has been proposed to guide the mismatch repair proteins to free termini, and therefore to recognize the newly replicated DNA strands [89].

The EXO1 gene was identified in S. cerevisiae as a gene encoding a double-stranded DNA-specific 5’ to 3’ exonuclease that interacts with MSH2 and functions in mismatch repair and genetic recombination [45,90]. It has recently been shown that yeast EXO1 interacts not only with MSH2 but also with MLH1 [91,92]. Similar observations were reported in human where EXO1 was found to bind strongly to hMSH2, hMSH3, and hMLH1 [93–95]. Yet, EXO1 mutants exhibit a weak mutator phenotype indicating that other exonucleases must function in eukaryotic MMR, just as in the case of E. coli where at least four exonucleases are involved in MMR [44,90–92]. Furthermore, it has been reported that the 3’ to 5’ proofreading exonuclease functions of the replication polymerases δ and ε participate in MMR [45,96]. As in E. coli, human MMR possesses a bidirectional excision capability which results in the excision of a large DNA fragment with an average size of 1 kb [97,98].

A human methyl-CpG-binding endonuclease, MED1, that has homology to bacterial base excision repair DNA N-glycosylases, forms a complex with hMLH1, and has been proposed to be the eukaryotic functional homolog of E. coli MutH [99,100]. However, although MED1 interacts with hMLH1 and binds to CpG-methylated DNA, human DNA mismatch repair operates independently of CpG methylation status, and there remains no evidence supporting a role for CpG hemimethylation as a strand discrimination signal [86]. Finally, the last steps of MMR consist in resynthesis of the excised strand, which requires the polymerases δ and possibly polymerase ε. The ligase that is involved in MMR has not been formally identified, but because DNA ligase I is often associated with polymerase δ and PCNA, it is a reasonable candidate [101,102].

4. Functions of the MutS- and MutL-related complexes other than MMR

In addition to post-replicative repair, MMR proteins have several other functions that are highly relevant to carcinogenesis. MMR components participate in various aspects of the recombination processes (reviewed in [5,11,13,29,103]). During homologous recombination, MMR plays a dual role in the processing of recombination intermediates containing mismatches. On the other hand, during gene conversion and single-strand annealing, the MMR proteins play an important role in removing non-homologous DNA [114–116]. The model proposes that MutSβ binds to the unpaired single-stranded DNA, stabilizes the recombination intermediates and allows the nucleotide excision repair Rad1–Rad10 cleave the 3’-ended tail [12,117]. Double-strand break repair through recombination requires Holliday junction resolution, a process which has recently been shown to be modulated by MMR [118–120]. Thus, the essential role of
MMR in double-strand break repair may explain why MMR-deficient cells are hypersensitive to ionizing radiation or topoisomerase inhibitors [16,24,25,121–123].

MMR components have been shown to participate in the recognition of DNA adducts caused by alkylating agents [18,124–137], platinum-based drugs including those used in cancer chemotherapy [138–144] or dietary agents, such as heterocyclic amines [145–147]. The recognition of the damaged bases by MMR initiates a signal transduction pathway that can trigger cell-cycle checkpoint and trigger apoptosis [16–22,122,130,147–149]. Both p53-dependent and p53-independent pathways participate in the MMR-dependent apoptotic response [17,18,21,130]. Some of the MMR components, hMSH2, hMSH6 and hMLH1, physically interact with BRCA1 in order to form a super complex that also contains BRCA2 and BLM, whose defects are associated with cancer predisposition [150–152]. It was recently reported that hMLH1 directly interacts with BLM and that this interaction is likely to be essential in some aspects of the recombination processes rather than for post-replicative repair [153]. The BASC complex has been proposed to recognize and repair aberrant DNA structures, participate in transcription-coupled repair of lesions induced by UV, ionizing radiation or reactive oxidative species, thus playing a key role in the genome stability.

5. Colorectal cancer susceptibility associated with MMR deficiency

The fact that some families appeared to have a predisposition to colorectal cancer (CRC) was first reported in 1913 by Dr. Warthin [154]. In 1966, Dr. Lynch studied two large families having multiple primary cancers, with the majority being adenocarcinomas of the colon, endometrium, and stomach, and demonstrated an autosomal dominant pattern of inheritance [155]. Some years later, the study of large kindreds using linkage analysis gave us insights into the etiology of CRC, in which two susceptibility loci could be mapped to chromosome 2p16 (which turned out to be 2p21–22) and 3p21 [156–159]. In order to identify the regions of loss of heterozygosity (LOH) in the tumors, the authors used a panel of polymorphic microsatellites that appeared to be unstable rather than lost. The microsatellite instability characteristic of these tumors was reminiscent of the mutator phenotype displayed by the MMR-deficient bacteria and yeast. This observation helped in identifying hMSH2, the human homolog of the MutS gene, on chromosome 2p [160,161], and hMLH1, the human homolog of E. coli MutL, on chromosome 3p [162]. Mutations of two additional MutL homologs, PMS1 and PMS2, were also found in patients with hereditary non-polyposis colon cancer [163].

No definite consensus criteria for defining the MSI* tumor phenotype have been defined yet [164,165]. The tumors are generally classified as MSI* tumors if instability is observed at a proportion of microsatellite loci, generally 30%, but some authors consider that MSI at a single locus, the BAT 26 mononucleotide marker, is enough to detect almost all MSI* tumors [164,165]. Moreover, some authors have tried to distinguish between tumors with instability at many loci, referred to as MSI-High, from those with few new alleles, defining the MSI-Low phenotype. While microsatellite instability is found in more than 90% of the HNPCC tumors, it is also seen in 15–20% of the sporadic colorectal cancers [166]. Moreover, MSI is not confined to the colon, but also affects the epithelial tumors of other organs including the stomach, endometrium, and ovary.

MSI* colorectal tumors are significantly more likely to be located proximal to the splenic flexure, to be poorly differentiated, to have a mucinous aspect, and to display peritumoral lymphocyte infiltration [167–169]. Furthermore, MSI* colorectal tumors are generally diploid and carry fewer p53 mutations [170–172]. In spite of the extensive studies, the clinical significance of the MSI* phenotype remains a matter of debate. Yet, most authors agree on the fact that MSI* colorectal tumors are associated with a better survival and a reduced likelihood of metastasis [167,168,173,174]. It is noteworthy that several studies have recently reported that the MSI* phenotype may influence the outcome of chemotherapy [24,175–177].

5.1. Hereditary nonpolyposis colorectal cancers

The HNPCC syndrome is characterized by an autosomal dominantly inherited predisposition to early-onset colorectal carcinoma and extracolonic epithelial-derived tumors most often located in the gastrointestinal and the urogenital tracts. The mean age of the HNPCC patients at diagnosis is 42 years instead of 65 years for sporadic CRC. HNPCC accounts for approximately 5% of all CRC, and is frequently associated with germline mutations of one of the mismatch repair genes, most often hMSH2 or hMLH1 [160–163]. The gene penetrance is very high, with a life-time risk for CRC that is over 80%, and greater for males and for hMSH2 mutation carriers [178–180]. Because CRC are among the most common cancers, the clinical selection criteria for families with high-risk CRC had to be defined for an efficient detection of germline mutation carriers. The criteria that were first established in Amsterdam in 1990 by the International Collaborative Group on HNPCC were too restrictive to allow effective diagnosis of HNPCC in small families, whereas large families with random clustering of CRC could be falsely diagnosed as HNPCC [181]. Thus, the criteria were modified in 1999 to take into account the occurrence of the extracolonic tumors [182,183]. The identification of germline mutations of either hMSH2 or hMLH1 could be performed in 50–70% of the families that met the Amsterdam criteria for HNPCC, whereas the families not complying with these criteria showed a much lower frequency of the MMR gene mutations [6,9,184–192]. Many studies have addressed the
question of the feasibility of molecular screening for HNPCC. So far, the detection of high-frequency MSI or the loss of either hMSH2 or hMLH1 immunostaining in the tumor, is the best indicator of the presence of a germline mutation in the high-risk patients with a family history [169,186,192–194]. The prevalence of germline mutations may represent an underestimate because some of the screening methods are unable to detect large deletions or promoter mutations. Yet, family linkage studies also report the involvement of the known MMR genes in approximately 60% of the families. These observations are in agreement with the prevalence of mutations identified in hMSH2 and hMLH1, and suggest the existence of additional genes in HNPCC.

5.1.1. Germline mutations of hMSH2

The hMSH2 gene consists of 15 exons, is located on chromosome 2p22–p21, and encodes a 934-amino acid protein (NCBI Ref. Seq. NM 000251 and NP 000242, for mRNA and protein, respectively). The mutations of hMSH2 represent approximately 40% of the mutations detected in HNPCC kindreds. The mutations of hMSH2 shown in Fig. 2 were assembled from the database of the International Collaborative Group on HNPCC (http://www.nfdht.nl/database/mdbchoice.htm) and the human gene mutation database (http://archive.uwcm.ac.uk/uwcm/mg/search/203983.html) [183,185,192,193,195–211]. Most of the mutations are unique, with the exception of an in-frame 3 bp deletion that has been found in seven unrelated families [196,202,204], and a splice defect due to a single-base substitution leading to the in-frame deletion of exon 5 that has been identified in 16 families of various geographical origins [195,197,199,201,202,204]. Among the 154 different hMSH2 mutations that have been analyzed, about 80% lead to premature termination of the hMSH2 gene product. Single-base substitutions directly resulting in a Stop codon account for one-fourth of the truncating mutations, whereas half are due to small insertions or deletions, causing frameshift generating downstream termination codons. Large genomic deletions, which account for the remaining truncating mutations, represent a frequent cause of HNPCC, with some of them showing inter-ethnic differences [208]. Concerning the missense mutations, we and others have attempted to distinguish between the neutral polymorphisms and the clinically relevant mutations by taking into account the following criteria: segregation in the family, inactivation of the wild-type allele in the tumor, absence of the variant sequence in control individuals (usually between 70 and 200), functional assays, and mapping of the variants to the equivalent residues in the bacterial MutS for which crystal structures are now available [32,33]. Some of the missense mutations of hMSH2 found in HNPCC patients have been examined in yeast strains in which the putative equivalent mutation has been introduced, and it was concluded that MSH2 mutations may result in a total or partial loss of function, and some mutations may exert a dominant-negative effect [212]. By contrast, some of the missense mutations located in the hMSH2 consensus interaction domain with hMSH3 or hMSH6 have no drastic effect [213].

5.1.2. Germline mutations of hMLH1

The hMLH1 locus is located on chromosome 3p21.3, consists of 19 exons, and encodes a 756-amino acid protein (NCBI Ref. Seq. NM 000249 and NP 000240, for mRNA and protein, respectively). Nearly 200 different mutations of hMLH1 have been described in the patients with HNPCC, representing 55% of the mutations reported so far. The mutations shown in Fig. 3 were assembled from the
While the mutations of hMSH2 were evenly distributed along the gene, a cluster of hMLH1 mutations has been found in the region spanning exons 15 and 16 [227]. Most of the hMLH1 mutations identified so far result in a truncated gene product by either frameshift or splice site mutations leading to exon deletion, which occurs at an unusually high incidence. The two ancestral founding mutations that account for the majority of the Finnish HNPCC patients are caused by the deletion of either exon 6 or exon 16, an event that has been proposed to result from Alu-mediated recombination [228]. Because MMR is highly conserved between yeast and human, various functional assays have been developed in yeast in order to distinguish between the non-pathogenic and pathogenic hMLH1 HNPCC mutations [229,230]. Based on the crystal structure of the 40 kDa ATPase fragment of the E. coli MutL, it has been proposed that missense mutations in the N-terminal conserved region of hMLH1 either affect ATP-binding and ATPase activity or alter the overall protein folding [36]. The recent elucidation of the X-ray crystal structure and function of the N-terminal 40 kDa fragment of hPMS2, which closely resembles hMLH1, also provided valuable information to elucidate the functional consequences of hMLH1 mutations [80]. The Thr117Met mutation, which has been found in nine unrelated families, maps to a conserved region that is supposed to mediate binding and/or hydrolysis of ATP [80]. Furthermore, the Thr117Met hMLH1 mutant is impaired in hMutLα formation and is unable to interact with hEXO1 [231]. The interacting domains of the three human MutL homologs have recently been determined: the C-terminal domain of hMLH1 interacts with 36 amino acid residues that are conserved within hPMS2, hMLH3 and hPMS1 [76,232]. A significant loss of hMLH1 interaction with hPMS2 has been observed for the HNPCC missense mutations located in the C-terminal region of hMLH1 [76,232]. Interestingly, the majority of the missense mutations that have been reported fall into either the ATPase domain or the MutL interacting domain (Fig. 3).

5.1.3. Germline mutations of hMSH6

The hMSH6 locus spans 27 kb on chromosome 2p16, consists of 11 exons, and encodes a 1360-amino acid protein (NCBI Ref. Seq. NM 000179 and NP 000170, for mRNA and protein, respectively). The first evidence that hMSH6 mutations could be involved in the development of colorectal cancer came from the description of mutations in two cell lines, HCT-15 and MT1 derived from tumors displaying microsatellite instability, primarily in mononucleotide repeats [233]. Germline mutations of hMSH6 were then reported in two atypical HNPCC Japanese families who lacked mutations in hMSH2 and hMLH1, with one family having a predominance of endometrial and ovarian carcinomas [234,235].

The frequency of hMSH6 germline mutations has been evaluated in various population-based series of colorectal cancer, including sporadic cases, familial non-HNPCC cases, as well as classical HNPCC cases, early-onset colorectal cancers and tumors with low levels of MSI [190,191,236–239]. Dominant mutations of hMSH6 have been identified and shown to cluster in two regions of the mutant protein: the N-terminal and the C-terminal, which are both MSH2 interaction domains and the ATP-binding site. By contrast, as shown in Table 3, the other germline missense mutations of hMSH6 are scattered along the gene. Additional functional studies are required in order to determine whether the missense mutations are neutral polymorphisms or clinically relevant mutations. Overall, these
studies concluded that germline mutations in hMSH6 are rare in all groups and confer predisposition to an atypical HNPCC phenotype, characterized by late-onset familial colorectal cancer, and frequent extracolonic tumors, particularly endometrial cancers [190,191,236–243]. Inactivation of hMSH6 causes a weaker mutator phenotype primarily confined to base substitutions [238,240]. A few frameshift mutations that are mainly restricted to insertion or deletion in mononucleotide tracts are also observed. Because the markers classically used to assess the MSI phenotype of tumors are predominantly dinucleotide repeats, the hMSH6-deficient tumors are generally classified in the group of MSI-Low tumors. The lower rate of microsatellite instability observed in hMSH6-deficient tumors is likely to reflect the partial functional redundancy of hMSSH6 and hMSH3 in mismatch repair.

5.1.4. Germline mutations of hMSH3

The hMSH3 gene locus spans 243 kb on chromosome 5q11–q12. It consists of 27 exons that encode a 1128-amino acid protein (NCBI Ref. Seq. NM 002439 and NP 002430, for mRNA and protein, respectively). The screening of 90 familial non-polypsis colon cancer without germline mutation of hMSH2 and hMLH1, did not allow the identification of any germline hMSH3 mutation [191]. In another study, 32 unselected MSI+ colorectal tumors were screened for germline mutations of hMSH3 [242]. Only one tumor harbored a germline missense variant changing Pro to Ser at codon 681, a residue which is conserved in the MSH3 mouse homolog, but replaced by a threonine in the yeast gene, an amino acid that is structurally similar to the variant serine. Thus, because this variant has not been detected in 190 individuals, it probably represents a rare polymorphism that lacks functional significance [242]. Thus, the role of hMSH3 in the initiation of colorectal carcinogenesis should be regarded as marginal.

5.1.5. Germline mutations of hPMS2

The PMS2 genomic locus encompasses 33 kb on chromosome 7p22, consists of 14 exons, and encodes an 862-amino acid protein (NCBI Ref. Seq. NM 000535 and NP 000526, for mRNA and protein, respectively) [244]. A non-expressed pseudogene has recently been described in this region of PMS2; because homology extends both in intronic and exonic sequences, polymorphisms in this pseudogene may be mistaken for mutations in the PMS2 gene [75]. This pseudogene may also represent a reservoir of mutations that could invade the functional hPMS2 gene by gene conversion or unequal crossing-over during meiosis or mitosis. In addition, a family of seven members of PMS2-related genes has been found in clusters on chromosome 7q [244]. Only one family with HNPPC has been reported to carry a pathogenic mutation of PMS2 that consisted of a 1230-bp genomic in-frame deletion resulting in the loss of codons 268–669, but evidence for segregation in the family was lacking [244]. The somatic inactivation of the wild-type allele in the tumor was an out-of-frame deletion which removed codons 301–381 [244]. The four additional germline mutations of PMS2 that have been reported, occur in patients with Turcot’s syndrome, a genetic disease characterized by an association of malignant brain tumors and CRC developing at a very young age. Two sisters inherited compound heterozygous frameshift mutations which both encoded truncated proteins that had lost the hMLH1 interaction domain [76,232,245]. One patient whose normal tissue displayed a high frequency of MSI had a missense mutation at codon 705, a region that has been involved in the interaction of hPMS2 with hMLH1 [232,246]. A nonsense mutation at codon 134 changing Arg to a Stop codon has been shown to confer a dominant mutator phenotype [247,248]. Several studies designed to better define the contribution of hPMS2 in hereditary colorectal cancer have searched for mutations in more than a hundred HNPPC and HNPPC-like families without known mutations in hMSH2, hMLH1, and hMSH6 [183,249,250]. No clear-cut pathogenic mutations of hPMS2 could be identified in these patients. Thus, the role of hPMS2 in familial colorectal cancer is highly questionable, except in the Turcot’s variant of the disease.

5.1.6. Germline mutations of hPMS1

Human PMS1 was identified on the basis of its homology with yeast PMS1 [163]. The hPMS1 locus consists of 13 exons spanning about 161 kb on chromosome 2q31–q33, and encodes a 932-amino acid protein (NCBI Ref. Seq. NM 000534 and NP 000525, for mRNA and protein, respectively). The only published hPMS1 mutation was a nonsense mutation, which resulted in exon-skipping in a patient with an MSI+ CRC. Although PMS1 is able to form a heterodimer with hMLH1, it is not involved in mismatch repair [82]. This observation is consistent with the fact that intestinal adenocarcinomas arising in Pms1-knockout mice do not display microsatellite instability [251]. Thus, these observations raised the possibility that the tumor from the patient with the hPMS1 germline mutation carried an undetected defect in another MMR gene. In fact, this case was recently reinvestigated and a germline mutation in hMSH2 was found, consistent with the MSI+ phenotype of the tumor. Moreover, another relative who shared the hMSH2 mutation, but not the hPMS1 mutation, developed CRC. Thus, predisposition to CRC in this family was linked to hMSH2, challenging the role of hPMS1 in colorectal carcinogenesis [250].

5.1.7. Germline mutations of hMLH3

In order to identify new genes that may function in mismatch repair, mammalian nuclear extracts were probed with the C-terminal interaction domain of hMLH1 using a far-western assay. This approach led to the identification of hMLH3, a 1429-amino acid protein of 160 kDa which is able to heterodimerize with hMLH1 and shares sequence homology with S. cerevisiae MLH3 [83]. Similarly to the
<table>
<thead>
<tr>
<th>Exon</th>
<th>Nucleotide change</th>
<th>Consequence</th>
<th>MSI phenotype</th>
<th>Functional analysis</th>
<th>Patient selection criteria</th>
<th>Tumor spectrum</th>
<th>Country origin</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>G → T at 431</td>
<td>Ser144Ile</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>Urogenital</td>
<td>The Netherlands</td>
<td>[237]</td>
</tr>
<tr>
<td>3</td>
<td>C → G at 467</td>
<td>Ser156Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Urogenital, Bladder</td>
<td>The Netherlands</td>
<td>[190]</td>
</tr>
<tr>
<td>4</td>
<td>Ins T at 650</td>
<td>FS, 217Stop</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>The Netherlands</td>
<td>[237]</td>
</tr>
<tr>
<td>4</td>
<td>Ins T at 650</td>
<td>FS, 217Stop</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>The Netherlands</td>
<td>[237]</td>
</tr>
<tr>
<td>4</td>
<td>Ins T at 652</td>
<td>Lys218Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC</td>
<td>The Netherlands</td>
<td>[239]</td>
</tr>
<tr>
<td>4</td>
<td>C → T at 742</td>
<td>Arg248Stop</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Breast, Endometrium, Bladder</td>
<td>The Netherlands</td>
<td>[190]</td>
</tr>
<tr>
<td>4</td>
<td>G → T at 854</td>
<td>Ser285Ile</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC</td>
<td>The Netherlands</td>
<td>[236]</td>
</tr>
<tr>
<td>5</td>
<td>T → C at 1019</td>
<td>Phe340Ser</td>
<td>Sporadic case</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>Not found in 199 controls</td>
<td>CRC, Breast, Leukemia</td>
<td>[239]</td>
</tr>
<tr>
<td>4</td>
<td>Del C at 1602</td>
<td>FS, 570Stop</td>
<td>MSI-High</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>Japan</td>
<td>[235]</td>
</tr>
<tr>
<td>4</td>
<td>G → C at 1636</td>
<td>Gln546Gln</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC</td>
<td>The Netherlands</td>
<td>[235]</td>
</tr>
<tr>
<td>4</td>
<td>G → A at 1696</td>
<td>G1668Arg</td>
<td>Partial loss of function</td>
<td>HNPCC</td>
<td>Not found in 80 controls</td>
<td>Familial</td>
<td>CRC</td>
<td>[236]</td>
</tr>
<tr>
<td>4</td>
<td>Del T at 1781</td>
<td>FS, 609Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>The Netherlands</td>
<td>[190]</td>
</tr>
<tr>
<td>4</td>
<td>Del TCAG at 2150</td>
<td>FS, 733Stop</td>
<td>Complete loss of function</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>Familial</td>
<td>CRC, Breast</td>
<td>[190]</td>
</tr>
<tr>
<td>4</td>
<td>C → T at 2191</td>
<td>Gln731Stop</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Breast, Endometrium, Bladder</td>
<td>Japan</td>
<td>[235]</td>
</tr>
<tr>
<td>4</td>
<td>A → G at 2408</td>
<td>Asp803Gly</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Lung</td>
<td>The Netherlands</td>
<td>[236]</td>
</tr>
<tr>
<td>4</td>
<td>TCT → C at 2674</td>
<td>FS, 891Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>The Netherlands</td>
<td>[190]</td>
</tr>
<tr>
<td>4</td>
<td>Del CT at 3052</td>
<td>FS, 1020Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Endometrium, Bladder</td>
<td>Finland/Sweden</td>
<td>[191]</td>
</tr>
<tr>
<td>5</td>
<td>C → A at 3259</td>
<td>Pro1087Thr</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Prostate, Lung</td>
<td>The Netherlands</td>
<td>[236]</td>
</tr>
<tr>
<td>4</td>
<td>Del C at 3262</td>
<td>FS, 1092Stop</td>
<td>MSI at MNR</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC, Stomach</td>
<td>Japan</td>
<td>[234]</td>
</tr>
<tr>
<td>4</td>
<td>Intron 5 a → g at 3439-2</td>
<td></td>
<td>Complete loss of function</td>
<td>HNPCC</td>
<td>Not found in 185 controls</td>
<td>Familial</td>
<td>CRC, Breast</td>
<td>[236]</td>
</tr>
<tr>
<td>4</td>
<td>Intron 7 a → c at 3647-2</td>
<td></td>
<td>Complete loss of function</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>Family</td>
<td>CRC, Endometrium, Breast</td>
<td>Denmark</td>
</tr>
<tr>
<td>5</td>
<td>C → T at 3772</td>
<td>Gln1258Stop</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC</td>
<td>The Netherlands</td>
<td>[237]</td>
</tr>
<tr>
<td>5</td>
<td>g → a at 4001+1</td>
<td>Splice defect</td>
<td>MSI-Low</td>
<td>HNPCC</td>
<td>No hMLH1 or hMSH2 mutation</td>
<td>CRC</td>
<td>The Netherlands</td>
<td>[190]</td>
</tr>
</tbody>
</table>

FS: frameshift mutation.
MNR: mononucleotide repeat.
MLH3-deficient yeast, cells that express a dominant-negative MLH3 mutation have been shown to display microsatellite instability [83,84]. Interestingly, hMLH3 gene maps to chromosome 1q42.43, a region showing frequent LOH in sporadic CRC, especially those of early-onset, and hMLH3 is highly expressed in intestinal epithelial cells [83,252]. Thus, it was tempting to propose MLH3 as a candidate gene for MSI+ familial CRC. A screening of germline hMLH3 mutations has been performed using single-strand conformation polymorphism (SSCP) analysis in 52 Finnish patients displaying features of inherited CRC. Although most of these tumors (88%) had been diagnosed as MSI+, no germline hMLH3 mutation could be found [253]. Using the denaturing gel gradient electrophoresis (DGGE) technique in order to screen 39 unrelated HNPCC families and 288 patients suspected to have HNPCC led to the identification of ten germline mutations, one frameshift and nine missense mutations in twelve patients with suspected HNPCC [254]. Three of these patients also carried a germline mutation in hMSH6, but did not seem to display a more severe phenotype than those with a single hMLH3 variant [254]. Yet, either segregation analyses or functional assays of the variants are needed to better define the possible role of hMLH3 in CRC. In a recent study, 60 probands with increased susceptibility to CRC but without identifiable germline mutations in hMSH2 or hMLH1, were compared with 24 MSI+ CRC tumors, and 12 CRC tumors without MSI, but with LOH at the hMLH3 locus (14q24) [255]. While no germline mutations could be identified in the former group of patients, one-fourth of the MSI+ tumors had accumulated somatic frameshift mutations in the coding polyadenine repeats of MLH3. Interestingly, two of the twelve tumors displaying 14q24 LOH had mutations leading to Stop codons, but whether these were germline or somatic was not reported [255]. In conclusion, future studies are required to definitely assess the role of hMLH3 in the progression of colorectal carcinogenesis.

5.1.8. Germline mutations of EXO1

Given that EXO1 functions in mismatch repair and is able to interact both with hMSH2 and with hMLH1, it was interesting to evaluate its possible involvement in familial CRC [93,95]. In a recent study, 33 of the HNPCC patients and the 225 patients suspected of HNPCC that were analyzed for hMLH3 mutations were screened for EXO1 germline mutations [256]. Fourteen mutations of EXO1 were detected, including one splice-site mutation in a family with HNPCC and 13 missense mutations in patients suspected of HNPCC. None of these tumors displayed mutations in hMSH2, hMLH1 or hMSH6, but the data concerning hMLH3 were not reported [256]. Although the pathogenic consequences of these mutations could not be definitely determined in the absence of a functional assay, none of these variants could be detected in more than 200 control individuals. Moreover, these changes were located in regions that are largely conserved, within domains necessary for the exonuclease activity or the interaction with hMSH2. Twelve of the tumors which arose in patients with a germline missense EXO1 mutation displayed LOH of the corresponding 1q42–43 region. Unexpectedly, these tumors retained the normal allele of EXO1 and lost the mutant allele, an observation that may indicate an haploinsufficiency effect of EXO1. In that respect, it is interesting to note that several missense mutations of hMLH1 found in HNPCC patients have been shown to disrupt not only hMutLα heterodimer assembly, but also hMLH1–hEXO1 complex formation [231].

In conclusion, germline involvement of MMR genes other than hMSH2 or hMLH1 in familial colorectal carcinomas is rare. Yet, it is most likely that other still unknown genes predispose to colorectal cancer in the families that are not attributable to mutations in hMSH2 or hMLH1. One may expect that the linkage studies that are currently extended to a genome-wide search will allow the identification of novel chromosomal regions and the further characterization of new genes associated with CRC [191]. Yet, this approach is difficult for the following reasons. First, because CRC are common cancers, it may be difficult to eliminate random clustering. Second, environmental factors are known to participate in colorectal carcinogenesis. Third, it is likely that the penetrance of the genes that remain to be discovered is lower than that of hMSH2 or hMLH1.

5.1.9. Inactivation of the remaining wild-type allele

Usually, HNPCC patients are heterozygous for a normal and a mutant allele of one of the MMR genes, most often hMSH2 and hMLH1. The DNA mismatch repair genes are believed to behave like tumor-suppressor genes in which two hits are required to cause a phenotypic effect. Apart from the very unusual case of EXO1, tumor development is linked to the inactivation of the wild-type allele, an event that probably occurs early in carcinogenesis [257,258]. The inactivation of the normal allele can occur by somatic deletion, point mutation or promoter methylation [259,260]. In HNPCC tumors carrying an hMSH2 mutation, somatic mutations are predominant upon LOH at the MSH2 locus, which occurs in approximately 10% of the tumors [259,261]. The hMLH1 promoter hypermethylation is the predominant mechanism of hMLH1 inactivation, contributing to approximately 50% of the cases [226]. LOH at the hMLH1 locus is more frequent than for hMSH2 and accounts for approximately one-third of the cases, whereas point mutations are rare events in hMLH1-associated HNPCC tumors [259,261]. LOH and hypermethylation are mutually exclusive mechanisms of hMLH1 inactivation [226]. Finally, it has been shown in yeast that strains heterozygous for some mutations have a mutator phenotype, as a result of either a dominant-negative effect or haploinsufficiency [212,229,230]. Thus, it is more than likely that some tumors have arisen in cells having both a wild-type and a mutated allele of MMR genes.
5.2. Sporadic MSI+ colorectal cancers

In sporadic MSI+ tumors, inactivation of hMLH1 accounts for more than 90% of cases [192-194,226-262-265]. Biallelic methylation of the hMLH1 promoter which results in gene silencing is the primary cause of loss of hMLH1 expression [226,264,266,267]. In some cases, hypermethylation is present without any apparent decrease in hMLH1 protein expression, suggesting that methylation may affect only one allele [226]. By contrast, the hMSH2 promoter has not been found to be prone to hypermethylation [264,267]. In sporadic MSI+ tumors, the rare mutations are more frequently detected in hMLH1 than in hMSH2 [226,263,268]. Allele loss occurs at both hMLH1 and hMSH2 loci in 20–40% of sporadic CRC with microsatellite instability [226,262,269,270]. LOH and/or somatic mutations may occur together with hMLH1 promoter methylation [226].

6. Microsatellite instability consequence: the functional inactivation of target genes

There is ample evidence that the repeated sequences, such as microsatellite, are highly susceptible to misalignment during replication, resulting in a 100-fold increase in their mutation rate. Numerous non-coding mono- and dinucleotide repeats are scattered in the genome, accumulate alterations in MMR-deficient tumors, and are commonly used to assess the MSI phenotype [164,165]. Repeated sequences, most often mononucleotide tracts, also exist in a number of human genes, which make them potential targets for frameshift mutations arising during replication, leading to truncated proteins [271,272]. In fact, a number of these genes have been shown to be altered in MSI+ colorectal tumors. They encode proteins involved in signal transduction (TGFβ-RII, IRAK, PTEN [273–276]), apoptosis and inflammation (BAX, caspase-5 [277,278]), transcription regulation (E2F4, TCF-4 [279,280]), and DNA repair (MSH6, MSH3, MLH3, MED-1, RAD50, DNA-PKcs, BLM [242,253,272,281–284]). The frequency of mutations varies depending on the gene that is assessed, and whether primary tumors or cell lines are studied [271,272,285–288]. For example, frameshift mutation within the coding poly-A10 tract of TGFβ-RII occurs at an extremely high-frequency (75–90%) in CRC with microsatellite instability from both HNPCC and sporadic patients, as well as in CRC cell lines. Insertion/deletion mutations in the mononucleotide repeats located within BAX (G8), TCF-4 (A9), IGFIIR (G8), and hMSH6 (C8) also occur at a significant rate in MSI+ colorectal tumors. Other genes, such as caspase-5 (A10), hMSH3 (A8), and RAD50 (A9) are less frequently inactivated in primary tumors, but show a high incidence of frameshift mutations in CRC cell lines [271,272].

Whether these genes are the true targets for inactivating frameshift mutations, playing a role in tumor initiation or progression, or only reflect the high mutation rate associated with MMR defect, is a major issue in the comprehension of the MSI-associated carcinogenesis. Actually, the identification of genes that play key roles in the multistep pathway of colorectal carcinogenesis is particularly relevant to the context of HNPCC where precursor lesions are not more frequent than in the general population, but more prone to malignant transformation as a result of the high mutation rate of MMR-deficient cells [289]. The importance of the nature of the gene that is targeted by microsatellite instability is well illustrated by the TGFβ-RII and APC genes.

The inactivation of TGFβ-RII, which is involved in the negative growth regulation of colonic epithelial cells by TGFβ, occurs early during the transition between colon adenoma and carcinoma [290,291]. Additional TGFβ-RII gene mutations located outside the poly-A10 coding repeat and inactivating its biological function have been reported in CRC [292,293]. Moreover, inactivation of TGFβ-RII frequently occurs in gastric MSI+ tumors, but is rare in endometrial MSI+ tumors [294]. Altogether, these observations support the idea that TGFβ-RII inactivation is a key step in human colorectal carcinogenesis. Unexpectedly, the majority of MSH2-knock out mice succumb to lymphoid tumors at an early age, rather than to colorectal carcinomas which only develop in older mice [111,112]. In order to define whether the TGFβ-RII gene contributes to the etiology and tissue-specificity of carcinogenesis associated with MMR-deficiency, we have compared the human and mouse TGFβ-RII sequences. As shown in Fig. 4A, the poly-A10 tract of the human TGFβ-RII gene is located in a conserved region, and encodes GluLysLysLys. The sequence of the corresponding region determined in ten different mouse strains was AAGAGAAAAG (unpublished data). The first A-to-G transition is silent, while the second A-to-G transition is conservative, changing Lys to Arg, two closely related amino acids. While the differences between the human and mouse TGFβ-RII proteins are little, those observed at the nucleotide level are likely to be significant. In fact, the repeated sequence present in the human gene is shortened and interrupted by a Gi in the mouse TGFβ-RII, which confers stability to this sequence by preventing it from polymerase slippage during replication. Thus, we propose that the low incidence of colonic tumors in MMR-deficient mice is due to the fact that, contrary to its human counterpart, the mouse TGFβ-RII gene is not a target for the MSI-associated mutagenesis.

The human APC gene, whose mutations cause familial adenomatous polyposis, provides another example of the key role played by a coding repeated sequence in carcinogenesis. A common polymorphism of the APC gene has been identified in 6% of the Ashkenazi Jews, and in 28% of individuals having a family history of CRC [295]. This polymorphism results from a T-to-A transversion at nucleotide 3920 of the APC gene and changes Ile to Lys at codon 1307. The Ile1307Lys variant allele substitutes an A8 repeat to the AAATAAAAA normally present at this position,
creating a sequence that is prone to polymerase slippage (Fig. 4B). Interestingly, analysis of the tumors occurring in individuals with the Ile1307Lys variant revealed the presence of somatic truncating APC mutations that exclusively arose from frameshift mutations in the A8 tract specific of this allele. Thus, as in the case of TGFβ-RII, cancer predisposition is caused by the presence in APC of a repeated sequence that is hypermutable, and becomes the target for somatic inactivation.

7. Conclusion

In conclusion, the inactivation of the mismatch repair either as a result of a germline mutation of one of the two major MMR genes, hMSH2 or hMLH1, or because of hypermethylation of the hMLH1 promoter, is closely associated with colorectal cancer predisposition. Nevertheless, the genetic background of a substantial number of HNPCC patients remains unexplained, indicating that other genes participate in MMR and play important roles in cancer susceptibility. Despite an increasing knowledge of the genetics and biochemistry of MMR, little is known about the factors that contribute to the etiology and tissue specificity of MSI-associated carcinogenesis. The identification of the tumor-suppressor genes containing hypermutable repeated sequences that make them potential targets for MSI-driven mutations should provide valuable insights into this issue. Finally, increasing evidence indicates that the prognosis and the chemosensitivity of the MSI colorectal tumors differ from those without instability. Thus, we believe that the MSI phenotype should be routinely assessed in order to improve the clinical management of patients with CRC.

Acknowledgements

We are very grateful to Pierre Netter, Alain Sarasin and Anne Stary for stimulating discussions and critical reading on this manuscript. We wish to take this opportunity to thank all our colleagues, Valérie Boige, Francesco Borrini, Brigitte Bressac-de-Palierets, Michel Dureux, Pierre Duvillard, David Fallik, Philippe Lasser, Catherine Miquel, Jean-Jacques Panthier, Filippo Rosselli, and Jean-Christophe Sabourin with whom we have shared ideas and discussions, and who contributed to our work.

Support was provided by grants from the ‘Institut Fédératif de Recherche 54, CNRS-Institut Gustave Roussy’ (#74538), the ‘Association pour la Recherche sur le Cancer’ (#5366), the ‘Ligue Nationale contre le Cancer-Comité de Seine-et-Marne’ and the ‘Groupement des Entreprises Françaises dans la Lutte contre le Cancer’.

References


H. Zhang, B. Richards, T. Wilson, M. Lloyd, A. Cranston, A. Thorburn, R. Fishel, M. Meuth, Apoptosis induced by overexpression of hMSH2 or hMLH1, Cancer Res. 59 (1999) 3021–3027.


D.L. Cooper, R.S. Lahue, P. Modrich, Methyl-directed mismatch repair is bidirectional, J. Biol. Chem. 268 (1993) 11823–11829.


W.M. Grady, A. Raiput, L. Myeroff, D.F. Liu, K. Kwon, J. Willis, S. Markowitz, Mutation of the type II transforming growth factor-beta receptor is coincident with the transformation of human colon adenomas to malignant carcinomas, Cancer Res. 58 (1998) 3101–3104.

