Microsatellite Instability in Colorectal Cancer

C. Richard Boland and Ajay Goel
GI Cancer Research Laboratory, Division of Gastroenterology, Department of Internal Medicine, Sammons Cancer Center and Baylor Research Institute, Baylor University Medical Center, Dallas, Texas

Abstract
Microsatellite instability (MSI) is a hypermutable phenotype caused by the loss of DNA mismatch repair activity. MSI is detected in about 15% of all colorectal cancers; 3% are of these are associated with Lynch syndrome and the other 12% are caused by sporadic, acquired hypermethylation of the promoter of the MLH1 gene, which occurs in tumors with the CpG island methylator phenotype. Colorectal tumors with MSI have distinctive features, including a tendency to arise in the proximal colon, lymphocytic infiltrate, and a poorly differentiated, mucinous or signet ring appearance. They have a slightly better prognosis than colorectal tumors without MSI and do not have the same response to chemotherapeutics. Discovery of MSI in colorectal tumors has increased awareness of the diversity of colorectal cancers and implications for specialized management of patients.

Keywords
Lynch Syndrome; DNA Mismatch Repair; Hereditary Nonpolyposis Colorectal Cancer

Cancers can be characterized by patterns of changes in gene expression. Genes that mediate tumorigenesis can be broadly characterized as oncogenes, which are activated by alterations; and tumor suppressor genes, which are inactivated during tumorigenesis. Oncogenes can encode growth factors or their receptors, signaling molecules, regulators of the cell cycle, and other factors that regulate cell proliferation and survival. Their oncogenicity can be induced by mutations that lead to overactive gene products, amplifications that alter copy number, alterations or rearrangements that affect promoter function, or modified interactions with regulators of transcription or epigenetic modification. Tumor suppressors restrain growth and proliferation, passage through the cell cycle, motility, invasion, or other functions related to stable differentiation. Genes that encode tumor suppressors are commonly inactivated by deletion, mutations, promoter methylation, or other changes in regulation. Colorectal cancers (CRCs) develop gradually over a long period of time through the sequential accumulation of genetic alterations that overcome the redundant control mechanisms built into each cell.1 Only a few mutations are common to most colorectal
tumors, but each tumor has a unique combination of genetic alterations. It is possible that no 2 CRCs are alike, a daunting consideration for the rational planning of treatment.

Models of CRC Pathogenesis

It was not immediately obvious that CRCs would be so diverse genetically. The initial attempt to characterize multistep carcinogenesis resulted in a novel conceptual model in which specific genetic alterations were associated with the sequential evolution of the neoplastic phenotype in the colon. In the model proposed by Fearon and Vogelstein, inactivation of the *adenomatous polyposis coli* (*APC*) gene led to the initial appearance of the adenoma from normal colonic mucosa, mutations in *KRAS* were associated with growth of the adenoma, genetic alterations (mainly deletions) to genes on chromosome 18q resulted in adenoma growth and progression, and biallelic loss or inactivation of *p53* mediated the adenoma-to-carcinoma transition. Chromosomal instability and loss of heterozygosity (LOH) are important steps in this process, and there is a relationship between the degree of LOH and tumor progression. Although this model, proposed in 1990, has been useful, only a few CRCs actually evolve along this pathway.

Fearon and Vogelstein were careful to characterize the genetic changes as “alterations” and did not presume to have determined the mechanisms involved. Some of the alterations were point mutations (as occurs in *KRAS*), whereas many were deletions of relatively large pieces of chromosome (such as on 18q, in which an arm or whole chromosome can be deleted). Point mutations can occur through a number of mechanisms, such as chemical carcinogenesis or unrepaired “natural” degradation of DNA. A diet-induced increase in proliferation can create replicative errors in DNA that exceed the ability of the cell to repair them. The origin of genetic deletions and LOH is complex and is addressed in the accompanying review of chromosomal instability.

A controversy emerged between investigators who thought that some type of genomic instability was required to produce all the mutations that are present in colorectal tumors vs investigators who reasoned that cells continuously generate a small number genetic alterations, and that mutations that provide proliferative or survival advantages are selected for during clonal expansion. It is not clear which model occurs during CRC pathogenesis, but at least 1 type of colorectal tumor is caused by the hypermutable phenotype known as micro-satellite instability (MSI).

MSI as a Unique Mechanism in Tumor Development

Shortly after publication of the multistep genetic model of colorectal carcinogenesis, many investigators began searching the genome for novel tumor suppressor genes using powerful techniques. In 1992, Manuel Perucho used an arbitrarily primed polymerase chain reaction (PCR), extracting DNA from colonic tissues and amplifying thousands of sequences using a small number of randomly chosen PCR primers. This technique yielded unique but reproducible electrophoretic signatures from each sample. Perucho’s group used arbitrarily primed PCR to amplify matched samples of DNA from colorectal tumors and adjacent normal colonic tissue, separated the PCR products electrophoretically, and compared the results to identify differences in PCR amplicons between normal and tumor tissues—DNA missing from the tumor would presumably represent somatic cell deletions that were candidate regions for new tumor suppressor genes.

Deleted DNA bands were found in most colorectal tumor samples. However, upon careful analysis of the comparative autoradiograms, the authors noted that 12% of ~130 tumors had bands that were not actually deleted, but were shorter in length and migrated slightly further down the gel. They analyzed the sequences in these bands and discovered that they
contained simple repetitive sequences (i.e., microsatellites), principally in polyadenine (A\textsubscript{n}) tracts associated with Alu sequences. The DNA obtained from this group of tumors had undergone somatic deletion of ≥1 of the adenine residues. The method was unbiased and quantitative, permitting the estimate that each tumor contained ~10\textsuperscript{5} such mutations. The tumors with these “ubiquitous somatic mutations at simple repeated sequences” had unique clinical and pathological characteristics. First, the tumors with these deletion mutations were significantly more likely to arise in the proximal colon, were less likely to be invasive, were less likely to have mutations in KRAS or p53, were more likely to appear poorly differentiated, and came from younger patients. Based upon these findings, the authors concluded that these deletions represented a unique pathway to tumor development and predicted that the “catastrophic loss of fidelity in the replication machinery of normal cells” that caused them might be hereditary, although they had no evidence for this.\textsuperscript{8}

Simultaneously, Stephen Thibodeau’s laboratory was studying dinucleotide repeat sequences (a type of micro-satellite) that were useful as tools for genetic mapping and analysis of LOH, and looking for novel tumor suppressor genes on chromosomes 5q, 15q, 17p, and 18q in colorectal tumors. They observed deletion mutations in the [CA\textsubscript{n}] sequences (called “CA repeats”) in these regions and coined the term microsatellite instability (which they termed MIN). They detected MSI in 25 of 90 CRCs (28%), and noted that these aberrations were heterogeneous in different tumors. They referred to a large deletion or expansion within the [CA\textsubscript{n}] tract as a “type I mutation,” and called a single 2-base-pair repeat change a minor or “type II mutation.” The implications of type I vs II mutations have never been elucidated, but they might reflect the cumulative nature of mutations in dinucleotide repeat sequences as they occur over time.\textsuperscript{9} Importantly, they found that 89% of tumors with MSI were in the proximal colon and that the patients with MSI CRCs had a better prognosis; like Perucho, Thibodeau et al recognized that this represented a unique pathway for tumor development that “does not involve loss of heterozygosity.”\textsuperscript{10}

Studies from an international consortium led by Bert Vogelstein in the United States and Albert de la Chapelle in Finland helped elucidate the clinical implications of MSI, using microsatellites (mainly, but not exclusively dinucleotide repeats) as tools for genome-wide linkage analysis of hereditary CRC. Aaltonen et al found significant linkage at chromosome 2p using the marker D2S123 in 2 large kindreds with Lynch syndrome (then called hereditary nonpolyposis colorectal cancer [HNPCC]).\textsuperscript{11} The authors used D2S123 to look for LOH at that locus in the CRC DNA, testing the hypothesis that this was the site of the tumor suppressor gene that caused hereditary CRC and that the second hit to the gene would be LOH. Instead, they observed MSI with the marker and then noted that deletion mutations in microsatellite sequences were widespread in hereditary CRCs. They referred to this as the replicative error phenotype. More importantly, they found MSI in 13% of sporadic CRC cases and correctly reasoned that the hereditary tumors and a subset of sporadic tumors shared a unique but common pathway of tumor development.\textsuperscript{12}

It became apparent that a subset of colorectal tumors (at least 12% and perhaps as much as 28%) were characterized by a large number of mutations at microsatellite sequences, formed in the proximal colon, had unique clinical features (such as a better prognosis than other types of CRC), and were hereditary. MSI was the first DNA marker available to identify hereditary CRC, although none of the investigators who discovered MSI in CRC understood mechanisms by which it developed.

A series of investigations led to the realization that MSI arises from defects in the DNA mismatch repair (MMR) system and the identification of the 4 genes that cause Lynch syndrome. Interestingly, microbial geneticists predicted that defects in the MMR system
were probably responsible for MSI. The timeframe of these discoveries are illustrated in Figure 1.

DNA MMR System

The “mutator phenotype” was discovered in bacteria in the 1970s and 1980s and was well-characterized by the beginning of the 1990s. It is caused by mutational inactivation of genes involved in DNA repair. The first human disease that was clearly associated with defects in DNA repair was xeroderma pigmentosa, a rare autosomal recessive disease caused by biallelic inactivating mutations in genes involved in nucleotide excision repair. Several other DNA repair systems had been characterized that included MMR, base excision repair, and a variety of nuclease and other DNA excision repair enzymes, yet no human diseases were associated with these.

In prokaryotes, the MMR system consists of a family of enzymes that detect S-phase DNA replication errors (i.e., those that result in mismatches between the 2 strands of DNA) in which the newly synthesized strand has incorporated the wrong nucleotide. DNA polymerase sometimes makes errors incorporating the correct number of bases during replication of long repetitive DNA sequences, such as microsatellites. Slippage during replication of a repetitive sequence creates a temporary insertion-deletion loop (IDL) that can be recognized and repaired by the MMR system, along with single base-pair mismatches. If these are not repaired, during the second round of replication the original parental strand is copied correctly, but the erroneously synthesized daughter strand (with the mismatch or IDL) contains a mutation. Single base-pair mismatches result in point mutations, whereas IDLs result in frame-shift mutations that usually lead to a downstream nonsense mutation; this results in production of a truncated, nonfunctional protein. This is the basis of MSI.

MMR is more complex in yeast and mammalian systems. The yeast homologues of the bacterial mutS and mutL genes were cloned and given the names Mut S homologue (MSH) and Mut L homologue (MLH). Then, additional homologous copies of these genes were cloned from yeast, giving rise to the terms MSH1 through MSH6, and MLH1 through MLH3. Another MutL homologue, called post-meiotic segregation-1 (PMS1), was also identified in yeast. Each of these genes had diverged from the initial mutS and mutL sequences and has been associated with specific repair functions in nuclear DNA, during meiosis, or in mitochondrial DNA. Phylogenetic analysis indicates that MSH1 was the founding member of the Mut S family and that the homologues diverged from this progenitor gene, but no highly homologous MSH1 gene is present in higher organisms.

More importantly, in the yeast and other eukaryotes, MutS and MutL proteins no longer function as homodimers. Instead, MSH2 forms a heterodimer with MSH6 or MSH3, giving rise to MutSa or MutSβ, respectively. These heterodimers have different relative abilities to bind to DNA mismatches and, as a result, yeast and other eukaryotes have a broader ability to recognize and repair different types of DNA misincorporation. For example, MutSa has a higher affinity for recognizing single base-pair mismatches. In fact, MSH6 was initially called the “GT-binding protein.” It was later discovered that the MSH3 gene encoded another MutS homologue that dimerized with MSH2 to create a complex with increased ability to bind to larger IDLs. The evolution of diverse homologues of mutS increased the cell’s ability to recognize and repair synthetic errors in DNA and increased replication fidelity in higher organisms. The phylogeny and biochemistry of factors involved in MMR have been reviewed.

Mammals have 4 homologues of the prokaryotic Mut L gene: MLH1, MLH3, PMS1, and PMS2. PMS1 was the first mutL homologue cloned in yeast; it was given a unique name
because of its functional role in meiosis. Tracing the genes phylogenetically is complicated because the yeast *PMS1, MLH1, MLH2,* and *MLH3* genes correspond to human *PMS2, MLH1, PMS1,* and *MLH3* genes, respectively. The function of the products of the _MutL_ homologues are not as clear as those of the _MutS_ homologues, but the encoded proteins function as heterodimers. _MLH1_ and _PMS2_ form _MutLa_, which mediates the interaction between the _MutS_ proteins and enzymes involved in long-patch excision in postreplication mismatch repair. However, _MLH1_ (the major _MutL_ homologue in humans) encodes a product that can dimerize with _PMS1_ to form _MutLβ_, which suppresses mutagenesis in yeast but has an uncertain function in humans. Moreover, the dimer of _MLH1_ and _MLH3_ (_MutLγ_) helps suppress IDL mutations and functions during meiosis in yeast, but its function in humans is unknown. Loss of _MLH1_ results in total loss of _MMR_ activity, but loss of _PMS2_ can be partially compensated by _MLH3_. It is not known how _PMS1_ fits into this model, but the _MLH1–PMS1_ heterodimer is not part of the canonical human _MMR_ system. This system is illustrated in the model in Figure 2.

**MSI is Caused by Deficiencies in MMR**

By 1993, it was recognized that about 15% of colorectal tumors have a unique mechanism of pathogenesis; they might have been first solid tumors with subclasses that had features so distinct they could be considered as separate diseases. To identify the different pathogenic mechanisms, researchers analyzed the PCR products of colorectal tumors with MSI.

Microsatellite sequences are abundant throughout the genome; they are polymorphic among individuals but are unique and uniform in length in every tissue in each person. The heterogeneity of dinucleotide repeats made them valuable for forensic, gene mapping, and allele discrimination analyses. Figure 3 illustrates amplification of dinucleotide and mononucleotide repeats in colonic tissues using autoradiography and gel electrophoresis.

The initial published examples of MSI were autoradiograms of radiolabeled PCR products; findings in yeast indicated that defects in genes that encode _MMR_ factors might be responsible for the MSI observed, although no human _MMR_ genes had been cloned.

**MMR Genes, HNPCC, and Lynch Syndrome**

**MSH2 and HNPCC**

Because much of the attention was focused on MSI in hereditary CRC, several groups set out to determine whether germline mutations in _MMR_ genes were responsible for Lynch syndrome. Fishel et al cloned the human _MSH2_ gene based on its homology to the yeast sequence, mapped it to human chromosome 2p22–21 (close to the locus implicated earlier in the year), and found a sequence variation in patients with familial CRC at the −6 position of intron 13 of _MSH2_. It was proposed that this variant altered splicing, inactivated the gene product, and was a germline mutation that mediated CRC pathogenesis. In fact, this is a common intronic polymorphism, but it was linked (in cis) to a true inactivating mutation that was found shortly thereafter. The association between _MSH2_ and hereditary CRC explained the MSI phenotype observed in colorectal tumor samples; _MSH2_ is located near the locus described in the first reported linkage study.

Leach et al confirmed the role of _MSH2_ mutations in hereditary CRC, identifying germline mutations in 3 kindreds with familial CRC: 1 leucine to proline missense mutation, 1 large in-frame deletion, and 1 nonsense mutation. These mutations were found in multiple affected members of the kindreds but not in individuals without cancer. Parsons et al showed that the CRC cell line, HCT116, which had MSI at [CA]n sequences, was deficient in _MMR_ activity. Within 7 months of the discovery of MSI in CRC, it was linked to a class of
genes that had not previously been of interest to cancer biologists. It was recognized that colorectal tumors might develop through loss of a DNA repair mechanism and that germline mutations in at least 1 of these genes mediated HNPCC.26

**MLH1, PMS2, PMS1, and MSH6**

Shortly after the 2p locus was linked with familial CRC, CRC in 3 Swedish families was linked to chromosome 3p21–23.27 The association between MSH2 and Lynch syndrome indicated that other genes that encode MMR factors might be located on 3p, which would represent a second familial CRC locus. Two groups found MLH1 on 3p21—Bronner et al identified the human homologue of yeast MLH1 and found germline mutations in multiple members of a CRC family with significant linkage to the 3p locus.28

At virtually the same time, the groups of Vogelstein and de la Chapelle scanned a DNA database for sequences homologous to the yeast MLH1, mapped one to chromosome 3p, and cloned the gene. Papadopoulos et al identified germline mutations in 7 members of 1 large Finnish CRC kindred that had been linked to 3p, found mutations in 3 more CRC kindreds, and reported a mutation in this gene in a CRC cell line.29 Using the same strategy of searching databases for homologous sequences, the human PMS1 and PMS2 genes were implicated in familial CRC shortly thereafter.30 The association between PMS2 and CRC has been established, but the role of PMS1 in sporadic or familial CRC is not clear. In the period from December 1993 to September 1994, 3 more genes were linked to familial CRC; it took another several years to identify the relationship between the MSH6 gene17 and Lynch syndrome.31 MSH6 is associated with somewhat atypical CRC kindreds with delayed onset of the cancer, which obscured the initial efforts at discovery. Once it was realized that MSH6 families did not fit the classical expectations of the Lynch Syndrome phenotype, the case for its involvement was secured.32

**Lynch Syndrome**

The focus of MSI in CRC was immediately aimed at Lynch syndrome because of its inherited and unique features.33 These patients develop tumors at early ages, often between 20 and 30 years old. They frequently have multiple tumors, including those of the colon, rectum, endometrium, stomach, ovary, urinary tract, small intestine, and other sites, but no increase in the frequency of cancers of the breast, lung, or prostate.34 Before 1993, there was no agreement on whether this syndrome actually represented a distinct disease.26

Families with clusters of CRC were therefore evaluated in linkage studies to identify loci of putative hereditary CRC genes. The Amsterdam Criteria were developed in 1991 so that research groups would have a uniform collection of families to study.35 These criteria called for 3 CRC cases in a family in which 1 individual was a first-degree relative of the other 2, CRC occurred in at least 2 generations (in which familial polyposis was excluded), and 1 affected family member was younger than 50 years of age. The Amsterdam II Criteria were developed in 1999 to acknowledge the involvement of noncolonic tumors in the phenotype.36

Studies of these kindreds facilitated the identification of genes responsible for this disease. Lynch syndrome is the hereditary disease caused by inactivating germline mutations in genes that encode MMR factors.37 There are 4 definite Lynch syndrome genes: MSH2, MLH1, MSH6, and PMS2.38 There are 2 genes that are candidates for Lynch syndrome (MLH3 and Exo1), but have been only rarely associated with familial CRC or the reported associations are believed to be coincidental. There has been no independent confirmation of a role for MLH3 in CRC and there has been a contradictory report on the
role of ExoI in this disease. However, both genes are an integral part of the MMR system, so there is rationale for considering these genes.

The Amsterdam Criteria are rigorous; it was expected that these would exclude many actual Lynch syndrome families because of the small sizes of families, missing information, and other practical issues. Although this was the case, it was surprising that 40%–60% of families that met the Amsterdam Criteria did not have a germline mutation in an MMR gene and the tumors did not have MSI. Families with non–Lynch syndrome CRC have been termed familial colorectal cancer, type X by 1 large collaboration. These families are characterized by microsatellite-stable tumors, a lower relative risk for CRC, the absence of excess tumors outside of the colon, and later onset of CRC. The genetic basis of this group is not known.

MMR Gene Mutations in Lynch Syndrome

To diagnose Lynch syndrome, MMR genes are analyzed, preferably using exon-by-exon sequencing with PCR primers that include relevant portions of the in-tron–exon boundaries; these detect point mutations and small insertion-deletion mutations that involve DNA sequences between the primers but not larger deletions or rearrangements, such as DNA sequence that are rejoined after a break, because the breakpoints are usually located deep within introns or completely outside of the coding regions. When DNA sequencing is performed, the tracing is examined for a double signal, ie, the simultaneous presence of a wild-type and mutant nucleotide in the same sequence position. When there is a deletion of an entire exon in which the breakpoints are outside of the primers for that exon, the analysis gives normal results because only 1 strand is sequenced, and conventional techniques cannot determine whether 1 or 2 alleles are detected. Large deletions can be detected by multiplex ligation-dependent probe amplification, which quantifies the number of alleles at each exon. Large deletions account for >33% of all MSH2 mutations associated with HNPCC; MSH2 is located within a dense cluster of Alu sequences, which are prone to internal recombination followed by excision of the DNA loop. Large rearrangements also occur with MLH1 and MSH6.

Analysis for mutations in DNA MMR genes does not end when a sequence variation is found; interpretation is another issue. Most deletions and nonsense mutations have pathologic consequences, but missense mutations are not always interpretable. It is important to determine how well the sequence has been preserved (indicating its functional significance) and if it is located in a region where the MMR proteins interact. Disruption of regions of protein–protein interaction by a nonconservative amino acid change could disrupt complex formation and function (see Figure 4). The mutant allele might need to be cloned and its product tested in a functional assay. A proportion of mutations identified in genetic testing are reported as “variants of uncertain significance,” and will remain so until we are better able to predict protein function based on its structure. It is interesting that there is a missense mutation in MLH1 (c415G→C; D132H) that affects adenosine triphosphate binding of the mutant gene product, increases risk for CRC, but does not always cause MSI in CRC cells.

MSI in Sporadic CRC

Two of the 3 initial descriptions of MSI were made in samples from sporadic colorectal tumors, rather than tumors from patients with familial CRC. Much of the subsequent attention to MSI was directed to familial CRC, but only about 3% of all CRCs come from Lynch syndrome families. Approximately 12%–17% of all colorectal tumors have MSI, depending upon the methods used to detect it; (references can be found in the Supplementary Material) most CRCs with MSI are sporadic.
The majority of colorectal tumors with MSI have lost expression of MLH1 and PMS2 protein. In 1997, Kane et al showed that MLH1 was silenced by methylation in most of these tumors. The characteristic features of sporadic CRC with MSI include the absence of significant familial clustering, biallelic methylation of the MLH1 promoter; absence of MLH1 and PMS2 proteins; and frequent mutation (usually V600E) in BRAF. The tumors are frequently diploid (74%) and patients with sporadic colorectal tumors with MSI have a better prognosis than those with non-MSI tumors. Whereas patients with Lynch syndrome patients are younger than those with sporadic CRC, those with sporadic CRCs with MSI are older. Loss of MLH1 expression increases with age; the gene is lost in ~50% of CRC patients older than 90 years of age.

**Sporadic MSI and CpG Island Methylator Phenotype**

MSI-associated sporadic CRCs arise through a process that involves the CpG island methylator phenotype (CIMP). About half of the genes in the human genome have promoters that are embedded in clusters of cytosine-guanosine residues called CpG islands. Cytosines in these regions can be methylated by DNA methyltransferases. Methylation is a means by which a cell permanently silences genes; it remains in the progeny of each cell. In vitro, it is possible to demethylate at least some promoters using the DNA methyltransferase inhibitor 5-azacytidine; however, upon removal of this agent, remethylation occurs. Histone modifications indicate which CpG sites are marked for permanent methylation.

Methylation increases with advancing age, is accelerated in the colon in response to chronic inflammation, and might be an adaptive response to injury. A subset of CRCs has been identified that is characterized by CIMP. These tumors progress by methylating tumor suppressor genes and have distinctive clinical features compared to non-CIMP colorectal tumors. Some of the commonly methylated promoters have no apparent role in tumor development and have been labeled “methylated in tumor” or MINT genes, but methylation also occurs in promoters of known tumor suppressor genes, such as p16 and insulin-like growth factor 2, and DNA repair genes such as methylguanine methyltransferase and MLH1.

The region of the MLH1 promoter in which methylation mediates gene silencing is the 3′ end, close to the start codon. The 5′ end of the promoter is also prone to methylation, but this is not of functional importance unless the methylation extends to the critical 3′ region. Therefore, specific CpG residues are more important than others in mediating gene silencing. Importantly, most of the CRCs with sporadic MSI come from a CIMP background, which creates an important distinction from Lynch syndrome tumors.

Consequently, it is essential to know more about a colorectal tumor than whether there is MSI or not. Lynch syndrome tumors are associated with germline mutations in DNA MMR genes, occur in younger individuals, can have KRAS mutations (but never BRAF mutations), and are associated with a better prognosis than non-MSI tumors. Most sporadic colorectal tumors with MSI occur in older individuals, have BRAF mutations in about half of the cases (V600E), are associated with a background of CIMP, and are associated with a reduced mortality. Curiously, BRAF mutations, which are detected in sporadic but not familial CRCs with MSI are associated with reduced mortality.

**Germline Epimutations**

To aide in the diagnosis of Lynch syndrome, researchers have searched for germline mutations in a coding regions, splice sites, or promoter regions of MMR genes. However, instead of a genetic mutation, some cases of Lynch syndrome are caused by epigenetic
inactivation of genes by promoter methylation. These events, originally described as germline epimutations, are better referred to as constitutional epimutations, to reflect the aberrant silencing of a gene that is normally active in somatic tissues in the absence of a sequence mutation. Constitutional epimutations of MLH1 and MSH2 have been identified in families with Lynch syndrome who have no apparent MMR gene sequence mutations. MSH2 methylation can arise from a deletion in the 3' end of the gene just upstream of MSH2 (TACSTD1)—the loss of the stop codon somehow results in methylation of the CpG island downstream, which leads to constitutional silencing of MSH2. Similar to germline mutations, constitutional epimutations in MSH2 demonstrate classic autosomal dominant inheritance and are associated with a 50% risk of transmission to the offspring.

**Pathophysiology of Colorectal Carcinogenesis with MSI**

CRCs with MSI were discovered, in part, because they were different from most other CRCs; the tumor cells had a tendency to be diploid and had less LOH. The mutational patterns were of interest because of the unique clinical features of these tumors. In 1995, Markowitz et al provided an important insight into the carcinogenetic pathway. Transforming growth factor-β (TGF-β) signaling inhibits proliferation in the colonic epithelium; Markowitz et al showed that the TGF-β type II receptor (TGFβR2) was not expressed in cell lines with MSI or in xenografts of these cells grown in nude mice, but was expressed by cell lines without MSI. TGFβR2 includes 2 microsatellite sequences. From nucleotides 1931 to 1936, there is a 6-base pair [GT]3 sequence (the inverse of a C–A repeat) and from nucleotides 709 to 718 there are 10 consecutive adenines (A10). Deletion mutations in these microsatellite sequences create frame shifts that inactivate the gene product. These were found predominantly in the MSI cell lines; cells with mutant forms of TGFβR2 did not slow proliferation in response to TGF-β. The most frequent mutation was a single base-pair deletion in the A10 sequence; Markowitz et al demonstrated that this mutation made the protein nonfunctional. Furthermore, the authors identified mutations in the A10 tract of 90% of 111 colorectal tumor samples found to have MSI, proving that this was a relevant site of this form of genomic instability.

Several other genes affected by MSI were then identified that encoded regulators of cell proliferation (GRB1, TCF-4, WISP3, activin receptor-2, insulin-like growth factor-2 receptor, axin-2, and CDX), the cell cycle or apoptosis (BAX, caspase-5, RIZ, BCL-10, PTEN, hG4-1, and FAS), and DNA repair (MBD-4, BLM, CHK1, MLH3, RAD50, MSH3, and MSH6) (see Table 1). Remarkably, every human MMR gene except MLH1 includes a mononucleotide repeat of at least A7, so the process of MMR could become increasingly defective with cumulative losses of components on the system. However, it is not clear how many of the mutations at these loci are of functional significance (which has been determined for TGFβR2), or whether some are simply markers of MSI, because biallelic inactivation of these genes has not been documented in all the tumors. The identification of the proximate targets of carcinogenesis in the setting of defective MMR activity indicates that MSI-associated colorectal tumorigenesis occurs through a different biological pathway compared to sporadic tumors. CRC cells with MSI can activate or inactivate the same signaling pathways as those without MSI, but different proteins within these pathways are involved. Genes are altered by different mechanisms in CRC cells with MSI compared to those without MSI. For example, sporadic colorectal tumors arise from a combination of mutations and LOH, resulting in biallelic inactivation of APC. Colorectal tumors with MSI have an increase in the number of point mutations compared to cancer cells without MSI, are more likely to be diploid, and do not have experienced
widespread allelic imbalance or LOH. A substantial portion of colorectal tumors with MSI have normal expression of APC, but have mutations in β-catenin that make it unable to interact with APC protein and undergo degradation,\textsuperscript{81,82} the biological equivalent of having no APC protein. Moreover, some colorectal tumors with MSI that have neither inactivated APC nor mutated β-catenin could have frame-shift mutations in factors further downstream in the WNT pathway, such as in the A\textsubscript{q} sequence of TCF-4.

Of equal importance are the differences in the pathogenesis of colorectal tumors with MSI that come from patients with Lynch syndrome compared with sporadic CRC cells with CIMP, as illustrated in Figure 5. For reasons that are not clear, \textit{BRAF} mutations are found in the majority of sporadic tumors with MSI but not in tumors that arise in patients with Lynch syndrome, which often have KRAS mutations.\textsuperscript{64,66} Activating mutations in \textit{BRAF} and \textit{KRAS} are functionally equivalent because their products function in series, in the mitogen-activated kinase signaling pathway, which regulates epithelial cell proliferation.\textsuperscript{83}

**MSI and Inflammatory Bowel Disease**

The risk of CRC is increased in patients with inflammatory bowel disease, but it is not completely clear how chronic inflammation mediates carcinogenesis. Inflammation potentially increases mutagenesis via the generation of oxidative stress and formation of free radicals; mucosal repair involves the stimulation of proliferation, which could expand populations of colon cells with mutations that provide proliferative or survival advantages. However, studies in 2 different in vitro models have shown that oxidative stress can “relax” DNA MMR activity,\textsuperscript{84,85} and another study has demonstrated an increase in mutation rates in cells exposed to oxidative stress.\textsuperscript{86} So, at the time of greatest mutational load, the DNA MMR system is underactive—a biological paradox. The same has been reported for base excision repair, which might also lead to MSI.\textsuperscript{87}

There is evidence for MSI in colon cells of patients with chronic ulcerative colitis. Using a panel of 5 dinucleotide repeat markers, Suzuki et al found that 21% of 63 colitis-associated tumors had at least 1 mutated microsatellite marker; a similar finding was reported for dysplastic lesions.\textsuperscript{88} Perhaps more surprisingly, Brentnall et al reported instability in at least 1 of 7 dinucleotide repeat markers in 50% of non-neoplastic mucosa samples from individuals with chronic ulcerative colitis, but not in samples from patients with acute infectious colitis;\textsuperscript{89} this finding was independently confirmed.\textsuperscript{90} These studies were done before the MSI marker panels were standardized, but an unpublished review of the primary data indicated that only 1 dinucleotide repeat was mutated in each instance, indicating low level MSI (MSI-low) rather than high MSI level (MSI-high).

**Models of MSI**

Studies of MMR expanded from \textit{Escherichia coli} and \textit{Saccharomyces cerevisiae} to human cells and diseases in the mid-1990s. Parsons et al discovered that HCT116 cells had defects in MMR activity and MSI\textsuperscript{25} that were associated with biallelic mutations in \textit{MLH1}. Stable transfer of human chromosome 3 into HCT116 (HCT116 + ch3 cells) restored \textit{MLH1} and created a model to study the role of MMR in the response to DNA damage and regulation of the cell cycle.\textsuperscript{91} MMR activity was restored in HCT116 + ch3 cells, which had reduced tolerance to DNA alkylation by \textit{N}-methyl-\textit{N}′-nitro-\textit{N}-nitrosoguanidine.\textsuperscript{91} The model also demonstrated that MMR regulates passage through the G\textsubscript{2}/M cell cycle checkpoint and the response to 6-thioguanine, which is incorporated into DNA as guanine but acts like a mismatch in newly synthesized DNA.\textsuperscript{92} MMR activity is required to restrain clonal expansion of cells exposed to \textit{N}-methyl-\textit{N}′-nitro-\textit{N}-nitrosoguanidine and prevent proliferation of cells with damaged DNA.\textsuperscript{93} Cultured CRC cells with intact MMR activity were significantly more sensitive to therapeutic concentrations of 5-fluorouracil (5-FU; 5–10
μM) than DNA MMR–deficient cells, as well as other chemotherapeutics. This provided the theoretical basis for later observations that patients with MSI CRCs might not respond to traditional 5-FU–based chemotherapy.

Cell models were subsequently developed to study other genes in the DNA MMR family. Although stable transfer of human chromosome 3 corrected some DNA MMR activity, HCT116 + ch3 cells were still devoid of MSH3 activity. Therefore, the HCT116 + 3 + 5 cell line was created by stable transfer of human chromosome 5 into HCT116 + ch3 cells, permitting an exploration of the effects of MSH3 on MMR. The CRC cell line HCT15 does not express MSH6 and the endometrial cancer cell line HEC59 does not express MSH2 or MSH6. Stable transfer of human chromosome 2 (which contains MSH2 and MSH6) into these cells created cell lines that permitted analysis of the role of the MutS system in human cells. In the modified HCT15 and HEC59 cell lines, MSH2 and MSH6 were re-expressed, MSI and the hyper-mutable phenotype were corrected (reducing mutations at the HPRT locus by 96% in HCT15), and the cells became more sensitive to cell death after exposure to mutagens. Similarly, in LoVo cells, which have defects at the MSH2 locus, MMR activity and sensitivity to mutagenic agents were restored following transfer of a fragment of chromosome 2.

Lynch syndrome is an autosomal dominant disease, associated with the inheritance of a single mutated gene. However, in cultured cells, defects in MMR are recessive; a single copy of the gene restores DNA MMR activity. MMR activity is lost in cells that lack MSH2 or MLH1, but only partially disrupted in cells without MSH6 or MSH3. These findings are important for understanding the pathogenesis of Lynch syndrome; patients with inactivating mutations in MSH6 develop an attenuated form of the disease and Lynch syndrome has not been associated with mutations in MSH3.

**Stoichiometry of MMR Proteins**

HCT116 (deficient in MLH1), LoVo (deficient in MSH2), HCT15 and DLD1 (deficient in MSH6), and SW48 (hypermethylated at MLH1) were used together with the single chromosome transfer-mediated restoration models and control lines (SW480, HEL, WI38) to study the stoichiometry of the members of the human DNA MMR system. MSH2 was the most abundantly expressed MMR protein; cells expressed 3–5-fold more molecules than its partner MLH1, depending upon the cell line studied. MSH6 protein was 4–12-fold more abundant than MSH3 protein and the combined amount of MSH6 and MSH3 expressed equaled that of MSH2 in cells. MSH6 might therefore have a more important role in DNA homeostasis than MSH3, and there are probably no other binding partners in the cell for MSH2. The relationship between expression levels of MSH2 and MLH1 have implications for MMR; a substantially larger number of MutS complexes are involved in identifying DNA mismatches than MutL complexes are required for interaction with the exonuclease complex. PMS2 was expressed in the lowest concentrations; the molar ratios of MLH1:PMS2 ranged from 1.4 to 2.75. Unfortunately, antibodies were not available to measure levels of PMS1 or MLH3 when these studies were performed. However, these studies helped provide a biochemical framework for the relative roles of MSH2 and MLH1 in MMR and their importance in human disease. Perhaps more importantly, these findings provided insight into the clinical differences between Lynch syndrome caused by mutations in MSH2 and MLH1, which have the most highly penetrant phenotypes, and the more attenuated phenotypes associated with germline mutations in MSH6 and PMS2.

**Methylation of MLH1**

CIMP and the acquired form of MSI induced by hypermethylation of the MLH1 locus were studied in the CRC cell lines SW48 and RKO. SW48 cells do not express MLH1 or PMS2.
protein, they lack MLH1 messenger RNA but do express stable PMS2 messenger RNA. Exposure of SW48 cells to the demethylating agent 5-aza-cytidine led to expression of the MLH1 messenger RNA and stable expression of the MLH1 and PMS2 proteins, which had been rapidly degraded in the absence of the stabilizing MLH1 protein. Demethylation-induced re-expression of MLH1 and PMS2 restored sensitivity to 5-FU, suggesting that demethylation might be used in conjunction with other therapeutic agents to overcome the drug resistance associated with acquired MSI in patients with CRC.

**Measuring MMR Deficiency**

Another model has been developed to measure mutation rates in cells with varying degrees of MMR deficiency. A green fluorescent protein reporter gene can be used to identify sites of microsatellite mutation as a surrogate for a direct measurement of MMR activity (which is labor intensive) to quantify the effects of physiological manipulations. This technology has shown that oxidative stress induces relaxation of MMR activity and increases frame-shift mutations. Furthermore, it has shown that mesalamine improves replication fidelity in cultured cells whereas aspirin does not.

There is no naturally occurring animal model of MMR deficiency, but knockout strains of mice have been created for each of the Lynch syndrome genes: MSH2, MLH1, PMS2, and MSH6. In spite of contrary claims, none of the heterozygous knockout models produces a phenotype that is similar to the human disease. The principal tumors that develop in these mice are lymphomas. Homozygous disruption of MMR genes produces animals with a constitutional deficiency of MMR activity, a variety of interesting genetic lesions, and some intestinal tumors, but are not useful in screens for agents that have therapeutic or preventive effects in humans. Animals with compound knockouts of Apc and a DNA MMR gene have accelerated polyp progression, but these tumors predominantly form in the small intestine.

The study of MMR in mouse models is limited because human genes associated with colorectal carcinogenesis include coding microsatellites, in places where the mouse genes do not. Furthermore, genes that are associated with colorectal tumorigenesis encode factors that are required for cell proliferation and survival, such as TGFβR2, BAX, and caspase-5, so it is a challenge to create mice with defects in these genes for cancer studies.

**MSI-High and MSI-Low**

At an international consensus meeting in 1997, the definition of MSI was standardized. The term MSI, when not otherwise modified, refers to MSH-high, in which >30% of the microsatellite marker panel is mutated, as defined in the Bethesda guidelines. The definition of MSI-high permits the characterization of a group of CRCs with similar clinical and pathological characteristics. However, if enough different microsatellite markers are used, a larger proportion of CRCs will be found to have at least 1 mutated microsatellite. Those CRCs in which at least 1 but <30% of the markers are mutated are called MSI-low and have clinical features of microsatellite-stable tumors. MSI-low has been observed in many tumor types.

Another type of MSI has been recognized that does not fit into the definition of MSI-high. This signature has been called “elevated microsatellite alterations at selected tetranucleotide repeats” (EMAST). EMAST is most frequently found in noncolonic tumors and is associated with p53 mutations. There is no evidence that it is caused by mutational inactivation of the MMR system.

Haugen et al have resolved the mechanistic basis of MSI-L and EMAST. They found evidence for EMAST in ~60% of sporadic CRCs and in all of the tumors that were MSI-
high or MSI-low. EMAST was also found in 35 of 82 (42.6%) of microsatellite-stable CRCs. Based on immunohistochemical analyses, 2%–50% of tumor cells were deficient in MSH3; its expression was heterogeneous throughout the tumor samples. Those parts of the tumor that did not express MSH3 were microdissected and found to be significantly more likely to have EMAST. Using in vitro models in which deficiencies in MMR can be studied in isolation, MSH3-deficient cells (including HCT116 + ch3, which has constitutional MSH3 deficiency, or cells in which MSH3 was conditionally knocked down), were found to have instability at di- and tetra-nucleotide repeats (ie, they had the EMAST and MSI-low phenotypes). This was important because MSI-low was detected through instability at dinucleotide repeat microsatellites. MSI-L and EMAST can therefore be induced by downregulation of MSH3, which occur heterogeneously (ie, non-clonally) throughout a tumor. It is possible that a similar mechanism is responsible for the appearance of MSI-low in the colonic mucosa of patients with chronic ulcerative colitis.

Diagnosis of MSI

Patients with colorectal tumors with MSI have longer survival times than patients with tumors without MSI. A study of 175 patients with Lynch syndrome (120 of whom had Lynch syndrome-MLH1 type), compared to more than 14,000 population-based patients with CRC, showed that the 5-year cumulative relative survival for patients with Lynch syndrome was 65%, compared to 44% of patients with sporadic CRC who were older than 65 years of age. A pooled analysis of MSI that included 32 studies and 7642 CRC cases found an overall hazard ratio of 0.65 for patients with tumors that have MSI. Patients with Lynch syndrome have lower-stage disease when they are identified compared to patients with other types of CRC, and it is less common for patients with Lynch syndrome to present with metastatic disease. Gryfe et al reported that 17% of 607 CRC patients who were younger than 50 years old had MSI; the hazard ratio for patients with MSI-associated tumors was 0.42, compared with patients in the same cohort with non-MSI tumors. Patients with tumors with MSI had lower mortality rates when patients were stratified by tumor stage, including patients with stage IV cancer.

Tumors with MSI have greater numbers of tumor-infiltrating lymphocytes that are activated and cytotoxic; the lymphocytic reaction is independently associated with longer survival. Studies report that colorectal tumors with lymphocytic infiltrates are associated with longer survival times; in fact, lymphocyte infiltration may be a better prognostic factor than routine pathology staging. Abnormal results from immunohistochemical analyses of MSH2 and MLH1 independently predict better outcomes in patients with stage II–T3 colorectal tumors. Therefore, detection of MSI in a patient with CRC is a positive prognostic factor, particularly among young patients.

However, not all studies have confirmed the value of MSI detection in prognosis. This might be because there are technical challenges in identifying tumors with MSI. Some studies could have included CRC cases that were falsely reported to have MSI-positive tumors, which might have altered the modest differences in predicted patient outcomes. Second, MSI analyses include 2 groups of patients: those with Lynch syndrome (who tend to be younger) and those with acquired methylation of the MLH1 gene (who are older). If Lynch syndrome patients are excluded from a cohort, the average age of patients with MSI is older—older patients have shorter overall survival times regardless of whether or not they have CRC. Most evidence, however, indicates that MSI predicts positive outcomes in patients with CRC.

The principal use of MSI testing in the clinic is to identify patients with Lynch syndrome. Approximately 15% of all colorectal tumors have MSI, and 75%–80% of this group have acquired methylation of MLH1; only ~2%–3% of all CRCs have germline mutations in one
of the MMR genes.\textsuperscript{49} Although MSI analysis is the first approach to identifying patients with Lynch syndrome, immunohistochemical analysis is equivalent in levels of sensitivity and specificity and is accessible to most pathology laboratories. Identification of abnormally expressed proteins can be used to identify genes that should be sequenced. MSI identifies MMR-deficient colorectal tumors with approximately 93\% sensitivity; most insensitivity is caused by mutations in \textit{MSH6}.\textsuperscript{128,129} Initial reports estimated that immunohistochemical analysis of MLH1 and MSH2 identified tumor samples with MMR defects with 92.3\% sensitivity and 100\% specificity.\textsuperscript{122} A comprehensive review of 16 studies that compared MSI with immunohistochemical analysis of MLH1 and MSH2 proteins suggested that immunohistochemistry was somewhat less sensitive overall, mainly because of its poor performance with MLH1. However, when expression of MSH6 and PMS2 were included in the analyses, the performance of immunohistochemistry improved, resolving the problems encountered with samples from patients with Lynch syndrome caused by MSH6 and PMS2.\textsuperscript{129} The limitation of immunohistochemistry is that staining can be heterogeneous throughout the tumor, which affects the sensitivity of the test.\textsuperscript{129}

Several panels of microsatellites have been used to diagnose MSI. The first consensus meeting recommended a panel of 3 dinucleotide repeats and 2 mono-nucleotide repeats that had been validated by a German consortium.\textsuperscript{110,111,130,131} This panel required that normal tissue be compared with tumor tissue, so many groups use a panel of 5 mononucleotide repeats that can be amplified and analyzed in a single assay. The location and nature of each of the microsatellite targets used for diagnostic purposes are listed in Table 2. The Pentaplex panel is as sensitive and specific as the initial panel proposed by the National Cancer Institute, but has some advantages that improve specificity. The \textsc{T25} repeat in the 3′ untranslated region of \textit{caspase-2} is useful in detecting loss of \textit{MSH6}.

### MSI as a Predictor of Response to Chemotherapy

Tests for MSI can be used to predict patient response to adjuvant chemotherapy, although this practice is controversial. MMR genes were identified in bacteria; inactivating mutations created a mutator phenotype and allowed tolerance to DNA-damaging agents.\textsuperscript{132} A similar phenotype was found in mammalian cell lines.\textsuperscript{16,133} The stable correction of MMR activity in cell lines restored the cytotoxic response to alkylating agents,\textsuperscript{91,93} 6-thio-guanine,\textsuperscript{92} 5-FU,\textsuperscript{94} platinum compounds, and other agents.\textsuperscript{96,134} This indicated that tumors with MSI might be resistant to some chemotherapeutic regimens.

However, the first study of the responsiveness of colorectal tumors with MSI reported just the opposite—tumors with MSI were more responsive to adjuvant chemotherapy than tumors without MSI.\textsuperscript{135} In this study, the assignment of patients to the treatment groups was flawed; instead of random assignment, oncologists selected patients for chemotherapy based on those with the fewest comorbidities and best performance status. Only 23 patients with MSI received chemotherapy, and the median age of these patients was 13 years younger than those who were not treated. Overall survival was significantly better in the MSI group, regardless of treatment. This created the impression that such patients were ideal candidates for adjuvant chemotherapy.

However, 11 subsequent studies have shown no benefit for chemotherapy among patients with colorectal tumors with MSI (see Table \textsuperscript{51,126,136 –143}). In the largest of these, a prospective, multicenter study in which patients were prospectively randomly assigned to treatment groups, the investigators found a nonsignificant, 2-fold excess in mortality among stage III CRC patients with MSI and a 3-fold increase in mortality among the stage II CRC patients with MSI.\textsuperscript{141} The MSI-associated increases in mortality might be related to the immunosuppressive effects of the treatment regimens, which would counter the increased anti-tumor immunity observed in patients with MSI.
A study by Fallik et al associated inclusion of the topoisomerase-I inhibitor, irinotecan, in the chemotherapeutic regimen with increased survival times of in patients with MSI-associated CRC,\textsuperscript{144} a finding that has been supported by Bertagnolli et al.\textsuperscript{140} It is possible that a combination of drugs without excessive immunotoxicity is best for patients with MSI-associated CRC. There is no evidence for the efficacy of traditional 5-FU–based adjuvant chemotherapy in patients with MSI-associated CRC, but there might be regimens that would benefit these patients. Patients with MSI-associated CRC should be included in clinical trials until the appropriate treatment regimen is identified. Also, these studies did not distinguish whether the colorectal tumors with MSI were sporadic or inherited. Because most tumors with MSI are sporadic, it is likely that these observations are more relevant to those tumors; additional studies focused on patients with Lynch syndrome will be needed to confirm this concept.

**Acknowledgments**

**Funding**

Supported by grants from the National Cancer Institute and National Institutes of Health R01 CA72851 (to Dr. Boland) and R01 CA129286 (to Drs. Goel and Boland).

**Abbreviations used in this paper**

- **APC**: adenomatous polyposis coli gene
- **CIMP**: CpG island methylator phenotype
- **CRC**: colorectal cancer
- **EMAST**: elevated microsatellite alterations at selected tetranucleotide repeats
- **5-FU**: 5-fluorouracil
- **HNPCC**: hereditary nonpolyposis colorectal cancer
- **IDL**: insertion-deletion loop
- **LOH**: loss of heterozygosity
- **MLH**: Mut L homologue
- **MMR**: mismatch repair
- **MSH**: Mut S homologue
- **MSI**: microsatellite instability
- **PCR**: polymerase chain reaction
- **PMS1**: post-meiotic segregation-1
- **TGF-\(\beta\)**: transforming growth factor-\(\beta\)
- **TGF-\(\beta\)R2**: TGF-\(\beta\) type II receptor

**References**


Gastroenterology. Author manuscript; available in PMC 2011 February 11.


Figure 1.
MSI research associated with colorectal cancer from 1990 to 2010. EpCAM, epithelial cell adhesion molecule.
Figure 2.
The DNA MMR system functions through a series of steps. (A) MSH2–MSH6 (MutSα) recognizes single base-pair mismatches, in which the DNA polymerase has matched the wrong base (G) with the T on the template (shown on left), and creates a sliding clamp around the DNA. This step that requires the exchange of adenosine triphosphate (ATP) for adenosine diphosphate (ADP) (by MSH2, but not MSH6 or MSH3). The complex diffuses away from the mismatch site, which is then bound by the MLH1-PMS2 (MutLα) complex (right). This “matchmaker” complex moves along the new DNA chain until it encounters the DNA polymerase complex. (B) The DNA MMR protein sliding clamp interacts with exonuclease-1, proliferating cell nuclear antigen (PCNA), and DNA polymerase. This complex excises the daughter strand back to the site of the mismatch (shown on left). Eventually, the complex falls off the DNA and resynthesis occurs, correcting the error. (C) Variations on the DNA MMR theme. Whereas MSH2–MSH6 recognizes single pair mismatches and small IDLs, MSH2–MSH3 (MutSβ) complements this by also recognizing larger IDLs (shown on left). The right side shows the possible interactions with different MutL dimers, as MLH1 can dimerize with PMS2, PMS1, or MLH3. The preferred interaction with MSH2–MSH3 is MLH1–MLH3 (MutLγ), but the precise roles of the other MutL heterodimers in this reaction are not entirely understood.
Figure 3.
MSI was initially discovered by autoradiography analyses of the PCR products after separation by gel electrophoresis (upper panel). N refers to DNA from the normal colon, and T from the tumor. The DNA polymerase used in PCR also has difficulty with the accurate amplification of templates, which is thought to be the explanation for the “ladder” of DNA bands that can be seen in the lanes for normal and tumor DNA. The upper panel illustrates the use of the 5 markers recommended by the National Cancer Institute consensus group; these consist of 3 dinucleotide repeats and 2 mononucleotide repeats (BAT25 and BAT26). In each instance, the DNA in the tumor has undergone somatic mutations (frequently, but not always, deletions), and the PCR product migrates to a different position on the gel, as indicated by the arrowheads. The lower panel shows the PCR products as they are analyzed by most laboratories using automated DNA sequencing with fluorescent primers. In this instance, 5 mononucleotide repeats have been analyzed, and in each instance, the mutations consist of deletions with different electrophoretic mobility (mutant alleles indicated by the arrows).
Figure 4.
The regions of protein–protein interactions among the members of the MMR system. Mutations in the regions of protein interactions are particularly likely to disrupt MMR function. Key interactions are illustrated among the MutS proteins, among the MutL proteins, and between these proteins and Exo1. ATP, adenosine triphosphate; PMS, post-meiotic segregation.
Figure 5.
There are 2 molecular pathways to the development of CRC with MSI. Approximately 20%–25% of colorectal tumors with MSI arise in individuals with Lynch syndrome. These tumors contain a germline mutation in one of the MMR genes, followed by a second hit to the wild-type copy (inherited from the unaffected parent); this could occur via LOH, methylation, or point mutation. Defects in MMR result in MSI and rapid accumulation of somatic mutations. It has been proposed that tumors arise via mutations in a few critical genes, but that large numbers of microsatellite mutations also occur—most of which are simply passengers that provide the mutational signature used to identify tumors with MSI. Colorectal tumors that arise in patients with Lynch syndrome often have mutations in KRAS. Most cases of CRC associated with MSI are not inherited (familial), but arise through sporadic methylation-induced silencing of MLH1. These sporadic tumors have the CIMP signature, resulting in methylation of many gene promoters. When the MLH1 promoter is methylated, MMR activity fails and MSI ensues. Thus, the mutational signature of sporadic colorectal tumors includes CIMP and MSI. BRAF mutations are also observed in most sporadic colorectal tumors, but do not occur in tumors of patients with Lynch syndrome.
### Table 1

The Genetic Targets of Microsatellite Instability in Colorectal Cancer

<table>
<thead>
<tr>
<th>Microsatellite length</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>A10</td>
<td>AIM2</td>
</tr>
<tr>
<td></td>
<td>CASPASE-5</td>
</tr>
<tr>
<td></td>
<td>MBD-4</td>
</tr>
<tr>
<td></td>
<td>OGT</td>
</tr>
<tr>
<td></td>
<td>SEC63 (also, A9)</td>
</tr>
<tr>
<td></td>
<td>TGFβ3, R2</td>
</tr>
<tr>
<td>A9</td>
<td>BLM</td>
</tr>
<tr>
<td></td>
<td>CHK1</td>
</tr>
<tr>
<td></td>
<td>GRB-14</td>
</tr>
<tr>
<td></td>
<td>MLH3</td>
</tr>
<tr>
<td></td>
<td>RAD50</td>
</tr>
<tr>
<td></td>
<td>RHAMM</td>
</tr>
<tr>
<td></td>
<td>RIZ (also, A8)</td>
</tr>
<tr>
<td></td>
<td>TCF-4</td>
</tr>
<tr>
<td></td>
<td>WISP3</td>
</tr>
<tr>
<td>A8</td>
<td>ACVRII</td>
</tr>
<tr>
<td></td>
<td>APAF</td>
</tr>
<tr>
<td></td>
<td>BCL-10</td>
</tr>
<tr>
<td></td>
<td>hG4-1</td>
</tr>
<tr>
<td></td>
<td>MSH3</td>
</tr>
<tr>
<td>A6</td>
<td>PTEN (2 A6's)</td>
</tr>
<tr>
<td>T10</td>
<td>OGT</td>
</tr>
<tr>
<td>T9</td>
<td>KIAA0971</td>
</tr>
<tr>
<td></td>
<td>NADH-UOB</td>
</tr>
<tr>
<td>G8</td>
<td>BAX</td>
</tr>
<tr>
<td></td>
<td>IGF2R</td>
</tr>
<tr>
<td>C9</td>
<td>SLC23A1</td>
</tr>
<tr>
<td>C8</td>
<td>MSH6</td>
</tr>
<tr>
<td>G7</td>
<td>AXIN-2 (A6, A6, C6)</td>
</tr>
<tr>
<td></td>
<td>CDX2</td>
</tr>
<tr>
<td>T7</td>
<td>FAS</td>
</tr>
</tbody>
</table>

Data from Duval and Hamelin.\(^78\)
### Table 2

Microsatellite Markers Used in Diagnosis of Microsatellite Instability in Colorectal Cancer

<table>
<thead>
<tr>
<th>Marker</th>
<th>Gene</th>
<th>MS repeat</th>
<th>Chromosomal location</th>
<th>Location of MS repeat</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2S123</td>
<td>Linked to MSH2</td>
<td>CA (n)</td>
<td>2p16</td>
<td></td>
</tr>
<tr>
<td>D5S346</td>
<td>Linked to APC</td>
<td>CA (n)</td>
<td>5q22-23</td>
<td></td>
</tr>
<tr>
<td>D17S250</td>
<td>Linked to p53</td>
<td>CA (n)</td>
<td>17q12</td>
<td></td>
</tr>
<tr>
<td>BAT-25</td>
<td>c-kit</td>
<td>A (25)</td>
<td>4q12</td>
<td>Intron 16</td>
</tr>
<tr>
<td>BAT-26</td>
<td>MSH2</td>
<td>A (26)</td>
<td>2p21</td>
<td>Intron 5</td>
</tr>
</tbody>
</table>

Pentaplex panel of mononucleotide repeat markers

| BAT-25 | c-kit                 | A (25)    | 4q12                 | Intron 16             |
| BAT-26 | MSH2                  | A (26)    | 2p21                 | Intron 5              |
| NR-21  | SLC7A8                | T (21)    | 14q11                | 5' UTR                |
| NR-24  | Zinc finger 2 (ZNF-2) | T (24)    | 2q11                 | 3' UTR                |
| NR-27  | Inhibitor of apoptosis protein-1 | A (27) | 11q22 | 5' UTR |

Other relevant microsatellite markers

| CAT25  | Caspase-2 (CASP2)     | T (25)    | 7q34                 | 3' UTR                |

MS, microsatellite; NCI, National Cancer Institute.
Table 3
Chemotherapy in Colorectal Cancer with Microsatellite Instability

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Study design</th>
<th>Adjuvant chemotherapy regimen</th>
<th>No. of patients (MSI/MSS)</th>
<th>Benefit of chemotherapy in patients with MSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elsaleh</td>
<td>2000</td>
<td>Consecutive patients</td>
<td>5-FU</td>
<td>63/669</td>
<td>Yes</td>
</tr>
<tr>
<td>Ribic</td>
<td>2003</td>
<td>Randomized controlled study</td>
<td>5-FU</td>
<td>95/475</td>
<td>No</td>
</tr>
<tr>
<td>Carethers</td>
<td>2004</td>
<td>Consecutive patients</td>
<td>5-FU</td>
<td>36/168</td>
<td>No</td>
</tr>
<tr>
<td>de Vos tot Nederveen Cappel</td>
<td>2004</td>
<td>Lynch syndrome patients</td>
<td>5-FU</td>
<td>28/0</td>
<td>No</td>
</tr>
<tr>
<td>Storojeva</td>
<td>2005</td>
<td>Randomized controlled study</td>
<td>5-FU/mitomycin</td>
<td>21/139</td>
<td>No</td>
</tr>
<tr>
<td>Benatti</td>
<td>2005</td>
<td>Consecutive patients</td>
<td>5-FU</td>
<td>256/1007</td>
<td>No</td>
</tr>
<tr>
<td>Popat</td>
<td>2005</td>
<td>Pooled data from multiple studies</td>
<td>5-FU</td>
<td>1277/6365</td>
<td>No</td>
</tr>
<tr>
<td>Lanza</td>
<td>2006</td>
<td>Consecutive patients</td>
<td>5-FU</td>
<td>75/288</td>
<td>No</td>
</tr>
<tr>
<td>Jover</td>
<td>2006</td>
<td>Consecutive patients</td>
<td>5-FU</td>
<td>66/888</td>
<td>No</td>
</tr>
<tr>
<td>Kim</td>
<td>2007</td>
<td>Prospective study</td>
<td>5-FU/leucovorin</td>
<td>98/444</td>
<td>No</td>
</tr>
<tr>
<td>Des Guetz</td>
<td>2009</td>
<td>Meta-analysis</td>
<td>—</td>
<td>454/2871</td>
<td>No</td>
</tr>
<tr>
<td>Bertagnoli</td>
<td>2009</td>
<td>Randomized controlled study</td>
<td>5-FU/irinotecan/leucovorin</td>
<td>106/677</td>
<td>No</td>
</tr>
</tbody>
</table>

5-FU, 5-fluorouracil; MSS, microsatellite stable.