We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Microsatellite Instability and its Significance to Hereditary and Sporadic Cancer

Jeffery W. Bacher, Linda Clipson, Leta S. Steffen and Richard B. Halberg

Abstract

Up to one million people within the United States may have Lynch syndrome (LS), but only 10% have been diagnosed. Early identification of these individuals is critical because they are predisposed to the development of colorectal and several other cancers at a relatively young age. Individuals with LS carry a germline mutation in one of four DNA mismatch repair genes, which leads to hypermutability in simple repetitive DNA sequences. This hallmark molecular phenotype called microsatellite instability (MSI) is now widely used to screen individuals needing germline sequencing to confirm diagnosis of LS. Standardized markers for MSI testing and other improvements in methodology have greatly improved the accuracy and cost-effectiveness of MSI testing. The current trend toward universal MSI screening of all colorectal and endometrial cancers will save lives by identifying LS prior to the development of deadly cancer. New technologies for MSI detection, such as next generation sequencing, open the possibility of a single test for LS that determines both tumor MSI status and germline mutations. Moreover, MSI detection is poised to take on an even greater role in prediction of responses to the new immunotherapies targeted at MSI-positive tumors.

Keywords: colon cancer, DNA mismatch repair, Lynch syndrome, microsatellite instability, MSI

1. Introduction

A form of hereditary colon cancer, we now call Lynch syndrome (LS), was first identified more than 100 years ago, but it was not until 1993 that rapid progress in unraveling the underlying genetic cause of this disease really began with the serendipitous discovery of a “mutator
phenotype” in colon cancers. The mutator phenotype observed in colon tumors was manifested as a high level of instability (i.e., insertion and deletion mutations) in simple repetitive sequences called microsatellites [1–3]. This form of genomic instability now referred to as microsatellite instability (MSI) has become the hallmark molecular signature of LS. Shortly after the discovery of MSI, the four DNA mismatch repair (MMR) genes responsible for LS were identified and the genetic basis for the disease was understood. The role of epigenetics in silencing the MMR system was later discovered, first in sporadic MSI cancers, then in LS. With this knowledge and the adoption of standardized guidelines for identifying and testing individuals at risk for LS, large scale screening for LS became possible and has set the stage for universal screening of all colorectal cancer (CRC) patients [4]. This milestone is important as the vast majority of individuals with LS are not diagnosed and an early detection of LS and identification of at-risk relatives is key to save lives. Finally, targeted immunotherapies offer new hope for treating the more challenging cases of hereditary and sporadic of MSI-positive CRC [5, 6].

2. Discovery of microsatellite instability (MSI) and its association with CRC

2.1. Microsatellite repeats

Microsatellite sequences are 1–6 base pair short tandem repeats that are highly mutable and ubiquitous in eukaryotic genomes [7–9]. As a consequence of high mutability, microsatellites tend to be quite variable in populations and therefore are widely used as molecular markers for linkage mapping, lineage mapping, and genotype identification purposes. Approximately 3% of the human genome contains microsatellite sequences, with mononucleotide repeats, predominantly poly (A/T) tracts, being the most abundant [10]. Microsatellite mutation rates vary greatly among loci, ranging from $\sim 10^{-6}$ to $\sim 10^{-2}$ mutations per locus per generation [11, 12]. The tendency of microsatellites to mutate increases with repeat number and can become pronounced beyond a critical number of repeats [10, 11, 13–15]. The vast majority of mutational variation can be attributed to intrinsic features of the locus, including repeat motif size, repeat number, and sequence composition. The major mechanism of mutagenesis in microsatellites is strand slippage during DNA replication, which can result in either insertion or deletion mutations in repetitive sequences, if not repaired effectively [16]. Post-replication, mismatch repair machinery removes any lesions occurring during replication to maintain genome stability.

2.2. Mutator phenotype hypothesis

In the early 1970s, Loeb [17–21] extended the concept of the “mutator phenotype” observed in bacteria to cancer biology. He proposed that high error rates due to alterations in DNA synthesis are causally linked to malignant transformation [18]. Loeb further speculated that high mutation rates caused by deficiencies in DNA repair activity could also contribute to cancer development. While earlier discoveries in bacteria had shown increased mutagenesis
due to defects in DNA polymerases and DNA repair, the contribution of Loeb was to propose a connection between a mutator phenotype and cancer development. The role of a mutator phenotype in cancer development is also integral to Nowell’s model [21] on tumor progression that is based on genomic instability providing the variability for clonal outgrowth and tumor evolution. The type of genomic instability that is described in this model is mainly chromosomal instability, in which breaks and rearrangements are increased as a consequence of inherited defects in DNA repair.

By 1991, Loeb [22] had refined his hypothesis arguing that an increased mutation rate or mutator phenotype could explain the high number of mutations believed to be present in many cancers that may be necessary for multistage tumor progression. He speculated that the spontaneous mutation rate in somatic cells is too low to account for the high number of mutations found in cancers and that an early step in tumorigenesis must be one that induces a mutator phenotype. Confirmation that the mutator phenotype contributes to at least some forms of CRC was conclusively demonstrated by the Cancer Genome Atlas Network study by measuring genome-wide mutation frequencies in 276 CRC samples [23]. Some (16%) of CRC samples were found to be hypermutated, with mutation frequencies 100-fold higher than nonhypermutated CRC. Interestingly, the hypermutated CRC tumors were found to have alterations in either DNA MMR genes or DNA polymerases.

2.3. Discovery of MSI in CRC

In 1990, Fearon and Vogelstein [24] published the multistep model of colon tumorigenesis in which they proposed that tumors develop as the result of mutational activation of oncogenes coupled with the mutational inactivation of tumor suppressor genes. Loss of a specific chromosomal region in CRCs was interpreted as evidence that the region contained a tumor suppressor gene and was detected as “loss of heterozygosity” (LOH) in a linked genetic marker. Following the publication of Fearon and Vogelstein, many investigators started looking for LOH events to determine the chromosomal location of potential tumor suppressor genes [24]. In 1993, Perucho and colleagues [1] performed arbitrarily primed polymerase chain reaction (PCR) to identify differences between normal and tumor samples from the colon. They noted that the amplicons actually became shorter in a few (12% of 130) tumors. Sequence analysis revealed that the PCR amplicons were composed of simple repetitive sequences, principally polyadenine tracts associated with Alu sequences in which one or more adenines were lost by somatic deletion in the cancers. These cancers, with an estimated $10^5$ ubiquitous somatic mutations in simple repetitive sequences, had unique clinical and pathological characteristics. First, these tumors were more likely to arise in the proximal colon, less likely to be invasive, less likely to harbor mutations in KRAS and TP53, and more likely to occur in younger patients. Based upon these findings, Perucho and colleagues [1] concluded that these tumors arose from a unique pathway involving the “catastrophic loss of fidelity in the replication machinery from normal cells” that caused them to be hereditary.

At the same time, two other groups were also using microsatellite markers to detect LOH to identify potential tumor suppressor genes [2, 3]. Thibodeau and colleagues [2] found that microsatellite repeats were often mutated in cancers, with alterations occurring in 25 out of 90
CRCs. They called this phenomenon “microsatellite instability” and used the abbreviation of “MIN.” The mutations were denoted as “type 1,” if the deletion or expansion was large and “type II,” if the change was limited to a single 2-basepair repeat change. The significance of this difference has never been fully resolved. CRCs with microsatellite instability were found to be primarily in the proximal colon and were associated with a better prognosis. Based on these findings, Thibodeau and colleagues [2] reasoned that this was a unique pathway to tumorigenesis that involved microsatellite instability and not chromosomal instability.

Another group, led by Vogelstein and de la Chapelle, was looking for LOH in LS families at the microsatellite marker D2S123, which they suspected was linked to a tumor suppressor gene causing hereditary CRC [3]. They observed many mutations at D2S123 and other microsatellites in the LS patients as well as 13% of sporadic CRC and called this phenomenon, replication error phenotype. Thus, three different groups had independently discovered MSI and named it either “ubiquitous somatic mutations in simple repetitive sequences,” microsatellite instability, or replication error phenotype. These names persisted until 2004 when the participants of the National Cancer Institute (NCI) workshop on MSI testing decided that the biomarker for identifying LS would be called microsatellite instability or MSI [25].

2.4. DNA mismatch repair systems

In the 1960s, 1970s, and early 1980s, laboratories studying bacteria [17, 26–28] and yeast [16] had discovered DNA mismatch repair and recognized that inactivation of the MMR genes resulted in widespread mutations at microsatellite sequences (i.e., a mutator phenotype). The first *Escherichia coli* mutator strain (*mutS1*) was isolated by Siegel and Byrson in 1967 and key MMR genes including *mutS, mutL, and mutH* were identified through genetic studies in the 1980s [17, 29, 30]. In vitro reconstitution of the *E. coli* MMR system from individual purified components facilitated mechanistic studies of individual *E. coli* MMR proteins [31, 32]. In the *E. coli* MMR system, a mismatched base is recognized by a MutS homodimer (Table 1). A MutL homodimer interacts with the MutS DNA complex, and then a MutH restriction endonuclease is activated by MutL. The MMR system recognizes the newly synthesized strand by the lack of methylation at GATC sites. MutH nicks the unmethylated error-containing strand to introduce an entry point for excision by helicases and exonucleases, and subsequent resynthesis by DNA polymerase III.

Shortly after the discovery of MSI in CRC, Strand and colleagues showed that MMR deficient mutants of the yeast strain *Saccharomyces cerevisiae* exhibited 100–700-fold increased instability in dinucleotide repeat tracts, demonstrating a clear link between loss of MMR and MSI [33]. The knowledge that instability in microsatellites was associated with loss of MMR activity, led to the rapid cloning and identification of the human homologs of yeast MMR genes [33–37]. Eukaryotic MMR systems were found to be more complex than in prokarroytes, but many features are conserved (Table 1) (recently reviewed by [38–40]). In eukaryotic MMR, MutS and MutL proteins do not function as homodimers, but instead form the heteroduplexes MSH2-MSH6 (MutSα) or MSH2-MSH3 (MutSβ) that bind to specific mismatches to initiate MMR. These heterodimers have different binding specificities, with MutSα being primarily responsible for repairing single base-base and insertion deletion
loop (IDL) mismatches, and MutSβ for repairing IDL mismatches. There are also multiple human homologs of the bacterial gene for MutL, including MLH1, MLH3, PMS1, and PMS2. Heterodimer MLH1-PMS2 (MutLα), the major MutL complex in humans, is involved in repairing a wide variety of mismatches. Two other MutL heterodimers, MLH1-PMS1 (MutLβ) and MLH1-MLH3 (MutLγ), appear to have a minor role in MMR. Proliferating cell nuclear antigen (PCNA) activates MutLα endonuclease activity to nick the DNA in a strand-specific fashion, which is then removed by EXO1 digestion and the new strand resynthesized by Polymerase δ and subsequently ligated [39, 41]. Mutator phenotypes conferred by defects in MSH3, PMS1, and MLH3 are much milder than those conferred by defects in MLH1, MSH2, MSH6, or PMS2, which are typically associated with LS.

<table>
<thead>
<tr>
<th>E. coli</th>
<th>S. cerevisiae</th>
<th>H. sapiens</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutS</td>
<td>MSH2-MSH6</td>
<td>MSH2-MSH6</td>
<td>Mismatch recognition, binds to single base and IDL mismatches</td>
</tr>
<tr>
<td></td>
<td>(MutSc)</td>
<td>(MutSc)</td>
<td></td>
</tr>
<tr>
<td>MutL</td>
<td>MLH1-MLH2</td>
<td>MLH1-MLH2</td>
<td>Strand incision, endonuclease activity</td>
</tr>
<tr>
<td></td>
<td>(MutLα)</td>
<td>(MutLα)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MLH1-MLH3</td>
<td>MLH1-MLH3</td>
<td>Strand incision, endonuclease activity</td>
</tr>
<tr>
<td></td>
<td>(MutLγ)</td>
<td>(MutLγ)</td>
<td></td>
</tr>
<tr>
<td>Dam methylase</td>
<td>Absent</td>
<td>Absent</td>
<td>Methylation of GATC sites in E. coli</td>
</tr>
<tr>
<td>MutH</td>
<td>Absent</td>
<td>Absent</td>
<td>Endonuclease nicks daughter strand at GATC sites, serves as strand discrimination signal in E. coli</td>
</tr>
<tr>
<td>RecJ, ExoI, ExoVII, ExoX</td>
<td>EXO1</td>
<td>EXO1</td>
<td>Strand excision, 5′-3′ dsDNA exonuclease</td>
</tr>
<tr>
<td>UrvD</td>
<td>None</td>
<td>None</td>
<td>Helicase, promotes strand excision</td>
</tr>
<tr>
<td>β-Clamp</td>
<td>PCNA</td>
<td>PCNA</td>
<td>DNA polymerase processivity factor</td>
</tr>
<tr>
<td>γ-Clamp</td>
<td>RFC</td>
<td>RFC</td>
<td>Loading of β-clamp/PCNA</td>
</tr>
<tr>
<td>SSB</td>
<td>RPA</td>
<td>RPA</td>
<td>ssDNA binding protein, acts in excision &amp; resynthesis</td>
</tr>
<tr>
<td>DNA Pol III</td>
<td>Pol delta</td>
<td>Pol delta</td>
<td>DNA polymerase involved in gap filling</td>
</tr>
<tr>
<td>DNA ligase</td>
<td>Unknown</td>
<td>Ligase I</td>
<td>Repair synthesis</td>
</tr>
</tbody>
</table>

Table 1. Mismatch repair proteins in E. coli, S. cerevisiae, and H. sapiens.

The rate of replication errors can vary by more than a million-fold, depending on the DNA polymerase and the local DNA sequence [39]. Correcting replication errors in MMR-defi-
cient and MMR-proficient cells can vary by more than 100,000-fold. The highest error rates in MMR-deficient yeast strains are for single-base insertion or deletion in long mononucleotide repetitive sequences, reflecting increased strand slippage during replication in these sequences. For example, Kunkel and Erie [39] reported the probability that a particular mismatch that will be made by a replicase varies from extremely rare misinsertion of the dCTP opposite template C by Pol α (≤10^−7) to much more frequent single-base deletion mismatches in long mononucleotide runs (≥10^−3). This high intrinsic error rate in replication of mononucleotide runs helps to explain why mononucleotide repeat markers are extremely sensitive to MSI in the absence of a functional MMR system.

2.5. MSI pathway in familial and sporadic CRC tumorigenesis

Many investigators support the view that some type of genomic instability is necessary to generate all the mutations observed in CRC, whereas others reason that mutations required to form cancer are accumulated spontaneously over long periods of time. Recent advances in molecular biology, especially sequencing, have revealed that CRCs are highly heterogeneous arising from several distinct pathways. Four types of genomic or epigenetic instability have been described in CRCs: chromosomal instability (CIN), microsatellite instability (MSI), CpG island methylator phenotype (CIMP), and global DNA hypomethylation.

About 3% of all MSI-positive CRCs are LS and about 15% are sporadic CRC [42, 43]. Tumor development in both LS and sporadic MSI-positive CRC involves the MSI pathway. The difference is that loss of MMR activity in LS tumors is the consequence of germline mutations or epimutations, while sporadic MSI-positive CRCs are caused by somatic methylation of the MLH1 promoter [44]. Sporadic CRCs with MSI are typically diploid, have biallelic methylation of the MLH1 promoter and subsequent loss of MLH1 protein expression, frequently have mutations in the BRAF gene, and are associated with better prognosis compared to individuals with non-MSI tumors. LS CRCs are also typically diploid and are associated with better prognosis, but have mutations in KRAS instead of BRAF, and have germline mutations or epimutations in MMR genes MLH1, MSH2, MSH6, and PMS2.

Tumorigenesis in MSI-positive CRC involves changes in the same signaling pathways as tumors without MSI, but often alterations occur in different genes and by different mechanisms. For example, initiating mutations in the APC gene are common in sporadic CRC. In contrast, a substantial portion of MSI-positive CRCs do not have mutations in the APC gene, but rather have mutations with similar consequences in CTNNB1 or other genes in the WNT signaling pathway. Sporadic non-MSI CRCs typically arise through CIN, whereas MSI-positive CRCs arise through the MSI pathway. The MSI pathway is characterized by a genome-wide increase in mutations, especially in microsatellite sequences. Since most microsatellites are in noncoding regions of the genome, mutations in these loci do not increase cancer risk. In contrast, mutations in short coding microsatellite sequences can lead to frame shift mutations and gene inactivation that are linked to cancer risk. For example, mutations in TGFβ-R2 occur primarily (>90%) in an A10 microsatellite tract that results in inactivation of the TGFβ-R2 protein [45, 46]. Transforming growth factor-β (TGF-β) signaling inhibits proliferation in the colonic epithelium and MSI-positive tumors often lack inhibitory TGF-β signaling due to
mutations in the gene for TGF-β type II receptor (TGFβ-R2). Loss of TGFβ signaling is a critical driver in the MSI pathway, but this is just the tip of the iceberg. There are an estimated 17,654 coding mononucleotide repeats in the human genome [47]. Sequencing of MSI-High (MSI-H) CRCs has identified recurring mutations in many other coding microsatellite sequences including tumor suppressor genes and DNA repair genes [48].

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>MSI-High, % Unselected</th>
<th>Cancer risk, % LS</th>
<th>References unselected; LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon</td>
<td>13%</td>
<td>10–80%</td>
<td>[58]; [49, 51, 59–62]</td>
</tr>
<tr>
<td>Endometrium</td>
<td>18–33%</td>
<td>15–71%</td>
<td>[63, 64]; [49–52, 59, 65, 66]</td>
</tr>
<tr>
<td>Stomach</td>
<td>22%</td>
<td>1–13%</td>
<td>[67]; [49, 53, 55, 59, 68–70]</td>
</tr>
<tr>
<td>Ovary</td>
<td>10%</td>
<td>4–20%</td>
<td>[71]; [49, 53, 54, 58, 59, 69, 72]</td>
</tr>
<tr>
<td>Small bowel</td>
<td>&lt;1–12%</td>
<td></td>
<td>[49, 53, 54, 59, 69]</td>
</tr>
<tr>
<td>Urinary tract</td>
<td>&lt;1–25%</td>
<td></td>
<td>[49, 53, 54, 69, 72, 73]</td>
</tr>
<tr>
<td>Skin (sebaceous tumors)</td>
<td>35–60%</td>
<td>1–9%</td>
<td>[74, 75]; [76–78]</td>
</tr>
<tr>
<td>Brain</td>
<td>1%</td>
<td>1–4%</td>
<td>[53, 69, 72, 79]</td>
</tr>
<tr>
<td>Prostate</td>
<td>1%</td>
<td>9–30%</td>
<td>[53, 54, 56, 57]</td>
</tr>
<tr>
<td>Breast</td>
<td>0–1%</td>
<td>5–18%</td>
<td>[80–83]; [49, 53–55, 84]</td>
</tr>
<tr>
<td>Hepatobiliary tract</td>
<td>16%</td>
<td>&lt;1–4%</td>
<td>[85]; [53, 58, 59, 68, 72]</td>
</tr>
<tr>
<td>Pancreas</td>
<td>1–4%</td>
<td></td>
<td>[53, 66, 86]</td>
</tr>
<tr>
<td>Thyroid</td>
<td>63%</td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>Skin (melanoma)</td>
<td>11%</td>
<td></td>
<td>[88]</td>
</tr>
<tr>
<td>Cervix</td>
<td>8%</td>
<td></td>
<td>[89]</td>
</tr>
<tr>
<td>Esophageal adenocarcinoma</td>
<td>7%</td>
<td></td>
<td>[90]</td>
</tr>
<tr>
<td>Sarcoma (soft tissue)</td>
<td>5%</td>
<td></td>
<td>[91]</td>
</tr>
</tbody>
</table>

1Percent of MSI-high tumors in unselected population.

2Cumulative risks of cancer by 70 years of age in germline MMR mutation carriers (for all MMR genes and both sexes) [49–91].

Table 2. Frequency of MSI cancers in LS and unselected populations.

Lifetime cancer risks for LS individuals vary depending upon which MMR gene is mutated and by gender. The majority (~80%) of LS tumors have mutations in MLH1 or MSH2. Lifetime risk of CRC by 70 years of age for MLH1 and MSH2 germline mutation carriers range from 40 to 80%, with higher risk for men (Table 2) [49]. The cumulative lifetime risk of CRC in MSH6 and PMS2 germline mutation carriers is lower, ranging from 10 to 22% [50–52]. LS individuals also have a significantly increased risk for a variety of extracolonic malignancies (Table 2). The highest risk is for endometrial cancer, which occurs in up to 71% of women with MSH6 mutations, 54% of those with MLH1 and MSH2 mutations, and 15% of those with PMS2 mutations [51, 52]. Significant cumulative risks also exist for cancers of the stomach, ovary,
small bowel, urinary tract, skin, pancreas, and brain. Breast and prostate cancers have not generally been considered part of the LS-associated cancer spectrum, but recent studies have found an increased risk for these cancers in germline MMR mutation carriers [53–57]. Dudley and colleagues [6] reviewed reports on the frequency of MSI-H across different tumor types in unselected populations, which includes both sporadic and familial cancers (Table 2).

3. Early development of MSI markers

3.1. Standardization of MSI testing

After the discovery of MSI in 1993, many laboratories began developing their own methods for measuring MSI and started to test different types of cancers. Unfortunately, there were no standards for MSI testing. Assays varied as to which and how many microsatellite markers to use. Moreover, investigators differed on the per cent of unstable markers necessary to classify a tumor as MSI-positive. This lack of standardization made it nearly impossible to compare results between laboratories and resulted in considerable variability in the frequency of MSI reported for a given tumor type. A number of studies were conducted to determine which microsatellite markers and what type of repeat motif was most sensitive and specific for the detection of MSI tumors. Two key studies described below provided the basis for markers chosen at the National Cancer Institute (NCI) workshop on MSI [25]. The first was by Dietmaier and colleagues [92] who tested 31 different microsatellites including six mononucleotide, 15 dinucleotide, three trinucleotide, five tetranucleotide, and two pentanucleotide repeats on a series of 58 primary CRCs. They found that sensitivity and specificity of markers were closely related to the type of the repeat (highest for mono and dinucleotide repeats) and that MSI could be subdivided into MSI-H (>20% of markers were unstable), MSI-Low (MSI-L) (<10% unstable markers), and microsatellite stable (0% unstable markers). The vast majority (14/15) of MSI-H tumors failed to express MSH2 or MLH1. In contrast, all of the MSI-L and MSI stable tumors had normal MMR expression. Based on these results, they recommended a diagnostic strategy for MSI assessment that utilizes a uniform panel of 10 microsatellites in which BAT26, BAT40, Mfd1S, D2S123, and D5S346 were tested first, followed by BAT25, D10S197, D18S58, D18S69, and MYCLJ if less than 40% of the initial set were mutated. Tumors were defined as MSI-positive if at least 40% of the tested markers were unstable.

The second study cited by the NCI workshop on MSI testing as a basis for the choice of MSI markers was a multicenter study to test the reliability and quality of MSI analysis [93]. Eight laboratories compared MSI analyses performed on 10 matched pairs of normal and tumor DNA from patients with CRC. They proposed that five microsatellite markers, which were selected from a panel of 30, should be analyzed in the first run and five additional microsatellite loci should be added in cases where less than two markers displayed MSI. A preferred set of five markers was not identified, but they suggested that the microsatellite panel should be comprised of different repeat types including mononucleotide and dinucleotide repeats. Cases with more than 40% unstable markers were classified as MSI-positive and those with less than 10% unstable markers were classified as MSI-negative [93].
In December 1997, the NCI sponsored an international workshop on Microsatellite Instability in Cancer Detection and Familial Predisposition to further review and unify the field [25]. The following recommendations (often referred to as the Bethesda guidelines) were made: (1) the form of genomic instability associated with defective MMR in tumors was to be called microsatellite instability or MSI, (2) a panel of five microsatellites (two mononucleotide repeats, BAT-26 and BAT-25; and three dinucleotide repeats, D5S346, D2S123, and D17S250) was recommended as a reference panel for MSI testing, (3) tumors should be classified as MSI-H if two or more of the five markers show instability, and MSI-L if only one of the five markers show instability, and MSI stable (MSS) if no markers were unstable, and (4) a unique clinical and pathological phenotype is identified for the MSI-H tumors, which comprise about 15% of colorectal cancers, whereas MSI-L and MSS tumors appear to be phenotypically similar. This standard was followed until 2004 when revisions were made at a second workshop.

The sensitivity, reproducibility, and cost effectiveness of MSI testing have improved considerably since the early days thanks to the use of all mononucleotide repeat markers and the introduction of fluorescent multiplex PCR and capillary electrophoresis technologies [94]. Currently, MSI testing involves comparing allelic patterns in microsatellite markers derived from a tumor and a normal (usually blood) samples from the same individual. A change in allele size between the normal and tumor samples indicates MSI. To generate the allelic profiles, DNA is extracted from each sample and amplified by PCR using fluorescently labeled primers flanking each microsatellite repeat locus. This is most efficiently done by multiplexing, allowing for simultaneous amplification and analysis of all markers in the panel. The resulting PCR products are resolved by capillary electrophoresis and the output is analyzed to determine allele sizes in comparison to known size standards [94]. The classification of tumor MSI status is based on the Bethesda guidelines [25].

3.2. Lynch syndrome screening guidelines

A number of different sets of criteria have been developed to identify patients who should be tested for LS (Box 1). The first set was the Amsterdam criteria in 1991, which was later modified to the Amsterdam II criteria in 1999 [95]. The Amsterdam criteria are very stringent and could miss as many as 58% of individuals with LS [96]. To address this limitation, the NCI published the Bethesda guidelines in 1997 and later the revised Bethesda guidelines in 2004 [97, 98]. Still, between 12 and 28% of individuals with LS could be missed using the revised Bethesda guidelines [4, 49]. To further increase sensitivity for the detection of LS, the trend has been

![Figure 1. National Comprehensive Cancer Network (NCCN) guidelines for LS testing.](https://dx.doi.org/10.5772/65065)
moving toward universal screening of all patients with newly diagnosed CRC. The National Comprehensive Cancer Network (NCCN) recommends either a selective approach using MSI/IHC to screen all patients with CRC diagnosed before 70 years of age and also those older patients who meet the Bethesda guidelines, or universal screening [99] (Figure 1). The selective strategy would miss only 4.9% of individuals with LS, whereas, universal screening would theoretically miss none, assuming 100% sensitivity [4].

Box 1. Lynch syndrome screening guidelines

Amsterdam criteria I (1991)

Three or more relatives with colorectal cancer, plus all of the following:

- One affected patient should be a first-degree relative of the other two
- Colorectal cancer should involve at least two generations
- At least one case of colorectal cancer should have been diagnosed before the age of 50 years

Amsterdam II criteria (1999)

Three or more relatives with LS-related cancer (colorectal cancer or cancer of the endometrium, small bowel, ureter, or renal pelvis) plus all of the following:

- One affected patient should be a first-degree relative of the other two
- Two or more successive generations should be affected
- Cancer in one or more affected relatives should be diagnosed before the age of 50 years
- Familial adenomatous polyposis should be excluded in any cases of colorectal cancer
- Tumors should be verified by pathological examination

Bethesda guidelines (1997)

Only one of the following criteria needs to be met:

- Cancer in families that fulfill the Amsterdam criteria
- Two LS-associated cancers in the same individual, including synchronous and metachronous CRC or associated extracolonic cancers (including endometrial, ovarian, gastric, hepatobiliary, or small-bowel cancer, or transitional-cell carcinoma of the renal pelvis or ureter)
- CRC and first-degree relative with CRC and/or LS-associated extracolonic cancers and/or colorectal adenoma; one of the cancers must have been diagnosed before the age of 45 years and the adenoma diagnosed before the age of 40 years
- CRC or endometrial cancer that was diagnosed before the age of 45 years
• Right-sided CRC with an undifferentiated pattern on histology, which is diagnosed before the age of 45 years
• Signet-ring-cell-type CRC that was diagnosed before the age of 45 years
• Adenoma that was diagnosed by the age of 40 years

Revised Bethesda guidelines (2003)

Only one of the following criteria needs to be met:
• CRC before the age of 50 years
• Synchronous or metachronous LS-related tumor
• CRC with 1 or more first-degree relatives with LS-related tumor before the age of 50 years
• CRC with 2 or more first- or second-degree relatives with LS-related tumor
• MSI in CRC in patient before the age of 60 years
• A panel of five quasi-monomorphic mononucleotide repeats may be more sensitive for MSI-High tumors than other microsatellite markers and may obviate the need for normal tissue for comparison

National Comprehensive Cancer Network (NCCN) guidelines (2015)

• Lynch syndrome tumor screening (i.e., MSI or IHC) should be performed for all patients with colorectal cancer diagnosed at or before the age of 70 years and also those after the age of 70 years who meet the Bethesda guidelines
• Or, universal MSI/IHC screening of all CRCs

4. Current use of microsatellite markers for detection of MSI

4.1. Mononucleotide repeats

In 2004, the revised Bethesda guidelines recommended the use of a panel of all mononucleotide repeat markers to increase the sensitivity of detection [98]. The recommendation was based on the observation that the original Bethesda MSI panel may underestimate the number of MSI-H tumors because of the use of dinucleotide repeats [100]. The revised guidelines indicate that the use of mononucleotide markers improves the sensitivity; hence, workshop participants suggested that more mononucleotide markers be used to evaluate MSI. The basis for the recommendation for the use of mononucleotide repeat markers is described below.

The BAT-26 mononucleotide repeat marker in the Bethesda panel is one of the most sensitive markers for MSI testing. Some investigators have suggested that MSI can be identified by analyzing tumor DNA with only BAT-26 [101, 102]. Zhou and colleagues analyzed 542 tumors
from various organs for MSI using a panel of 10 or more microsatellite markers versus BAT-26 [103]. They found concordance of results in 539 out of 542 (99.5%) cases [103].

An unusual property of BAT-26 and a few other microsatellite markers is that most individuals in the population have a single allele and are thus quasi-monomorphic, which permits MSI testing using tumor samples only [101, 102]. However, others have shown germline polymorphisms in BAT-26, especially in certain racial groups. For example, a study by Samowitz and colleagues found 7.7% of African Americans are polymorphic for BAT-26 [104]. A more extensive population study performed by Bacher and colleagues, which included individuals of Caucasian, African, and Asian descent, found low-level germline variation in BAT-26 and other quasi-monomorphic markers (Table 3) [94]. Thus, polymorphisms in these microsatellites limit their utility in MSI determinations without the corresponding normal DNA.

<table>
<thead>
<tr>
<th></th>
<th>NR-21 (%)</th>
<th>NR-24 (%)</th>
<th>BAT-25 (%)</th>
<th>BAT-26 (%)</th>
<th>MONO-27 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian-American</td>
<td>1.1</td>
<td>0.0</td>
<td>0.5</td>
<td>0.6</td>
<td>0.0</td>
</tr>
<tr>
<td>African-American</td>
<td>0.8</td>
<td>0.9</td>
<td>9.9</td>
<td>9.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Asian-American</td>
<td>5.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 3. Quasi-monomorphic mononucleotide repeat markers (% polymorphic alleles).

Inclusion of dinucleotide repeats in the Bethesda panel might lead to misclassification of some cancers. Incorrect assignments can result from a number of different factors. First, dinucleotide repeats are less sensitive to MSI than mononucleotide repeats [94, 105]. Second, instability involving only dinucleotide markers can occur in MSS tumors [94, 101, 106]. Third, size alterations in dinucleotide repeats can be difficult to interpret. Finally, mutations in MSH6 often do not lead to alterations in dinucleotide repeats [107]. These limitations lead Suraweera and colleagues [100] to propose using a panel of five quasi-monomorphic mononucleotide repeats (BAT-25, BAT-26, NR-21, NR-22 and NR-24). They determined the MSI status of 124 colon tumors, 50 gastric tumors, 20 endometrial tumors, and 16 colon cancer cell lines that had been previously established. The results were 100% concordant.

![Figure 2. The relative sensitivity and specificity of mono-, di-, tetra- and penta-nucleotide repeats for detection of MSI in mismatch repair deficient CRC.](image-url)
To determine the best markers for the MSI testing, a study of 266 mono-, di-, tetra-, and pentanucleotide repeat markers was conducted to identify those with the highest sensitivity and specificity for the detection of MSI in MMR deficient tumors [94]. A subset of each marker type was used to screen 225 human colon tumor samples that had been previously characterized for mismatch repair status. Consistent with previous studies, mononucleotide repeats were found to be the most sensitive and specific type of microsatellite marker for the detection of MSI (Figure 2). Based on this study, the MSI Analysis System (Promega Corporation, Madison, United States) was developed; it contains five quasi-monomorphic mononucleotide repeats, BAT-25, BAT-26, NR-21, NR-24, and MONO-27. The MSI Analysis System has several advantages over the Bethesda panel, including: (1) increased sensitivity and specificity, (2) easier interpretation of MSI patterns in mononucleotide repeats compared to dinucleotide repeats, (3) the quasi-monomorphic nature of the markers simplifies analysis and allows MSI classification in cases where only tumor samples are available, and (4) the inclusion of two highly polymorphic pentanucleotide repeats to prevent sample mix-ups (Figure 3) [94, 108]. This MSI kit is now a widely used alternative to the Bethesda panel [94, 108, 109].

Figure 3. MSI analysis of CRC and the corresponding normal sample using the MSI Analysis System (Promega Corporation, Madison, United States). The electropherogram shows the allelic profile generated from a normal sample (top panel) and a matching MLH1-deficient tumor sample (bottom panel). New alleles in tumor sample that are not found in matching normal are indicated by arrows. The panel contains five quasi-monomorphic mononucleotides for MSI determinations, including: JOE-labeled (green) NR-21, BAT-25 and MONO-27; fluorescein-labeled (blue) BAT-26 and TMR-labeled (black) NR-24. Two highly polymorphic pentanucleotide repeats (Penta-C and Penta-D) are included for sample identification in case of sample mix-ups or contamination problems.

4.2. Relative utility of MSI and IHC for Lynch syndrome screening

Commonly used screening tools for LS include: family history, tumor pathology, MSI, and MMR protein detection by immunohistochemistry (IHC). It has been found that family history and tumor pathology lack sensitivity and specificity for selecting patients for germline
mutation analysis [4]. In contrast, both MSI and IHC are highly effective strategies. Which method to use as the primary screening method for the detection of LS is a subject of ongoing debate [110, 111].

As the hallmark molecular signature of LS, MSI is widely accepted as a primary method for identifying individuals at risk for LS. Recent improvements in MSI testing have significantly enhanced accuracy and reduced cost. The advantages of MSI as a screening method for LS include: (1) high sensitivity for the detection of MMR loss, (2) use of quasi-monomorphic mononucleotide repeat markers which simplifies data interpretation and allows analysis of tumor samples alone when matching normal is not available, (3) utilization of fluorescent multiplex PCR technology that reduces labor, time, and cost of testing, (4) relatively easy interpretation, and (5) excellent intra- and interlaboratory reproducibility. Disadvantages of MSI testing include: (1) lack of specificity for LS as sporadic MSI is common, (2) failure to identify which MMR gene is involved, and (3) 5–10% false negative rate.

Advantages of IHC testing include: (1) high sensitivity and specificity for MMR loss, (2) wide availability in general pathology laboratories, and (3) identification of which MMR gene is mutated. IHC has several disadvantages including: (1) requirement for an experienced pathologist to interpret the results, (2) variable staining pattern, resulting in uncertainty in interpretation, (3) dependence of sensitivity on the antibody panel used, (4) possible lack of reliability in small biopsy samples, and (5) potential loss of antigenicity owing to nonpathogenic mutations, which can lead to a 5–10% false negative rate.

The significance, use and implications for MSI and IHC testing are similar, although the tests are slightly complementary. NCCN guidelines state that both MSI and IHC miss about 5–10% of cases [99]. Therefore, many labs have adopted the practice of using both MSI and IHC to maximize sensitivity for the detection of LS.

5. Emerging applications for MSI testing

5.1. Universal screening for Lynch syndrome

Up to one million individuals within the United States may have LS, but less than 5–10% are likely to have been diagnosed [112, 113]. The optimal strategy for identifying individuals with LS is a subject of continued debate. Some advocate targeted screening based on age of onset, family history, and/or histologic criteria to reduce the number of unnecessary tests. Others prefer universal screening of all CRCs to maximize sensitivity and improve outcomes through early monitoring. For example, Moreira and colleagues compared various strategies for identifying patients with LS and found that the revised Bethesda guidelines had a sensitivity of 87.8% compared with 100% sensitivity of the universal screening approach [4].

To help identify the undiagnosed cases of LS, the NCCN recommends that institutions use either a selective approach of testing all patients with CRC diagnosed before 70 years of age plus those diagnosed at older ages who meet the Bethesda Criteria, or universal testing. Universal MSI/IHC testing on all newly diagnosed colorectal and endometrial cancers
regardless of family history is practiced by many NCCN member institutions and other comprehensive cancer centers to identify which patients should have genetic testing for LS [114–117]. Universal screening has been shown to be cost effective for colorectal cancers and is endorsed by the Evaluation of Genomic Applications in Practice and Prevention working group at the Centers for Disease Control and Prevention (CDC), the US Multi-society Task Force on Colorectal Cancer, and the European Society of Medical Oncology [118–121]. The Cleveland Clinic has implementing universal MSI/IHC screening since 2004 [122]. Similarly, Ohio State University Comprehensive Cancer Center has screened all CRC patients for LS since 2006 and projects that if universal screening were adopted nationwide it could save thousands of lives every year (Figure 4) [112].

![Figure 4](image.png)

Figure 4. Early identification of Lynch syndrome patients saves lives. In 2014, 4100 or 3% of patients newly diagnosed with CRC carry germline mutations in MMR genes. Each of these individuals will on average have three at-risk relatives. Thus, 16,400 people per year could be potentially saved if all LS patients were identified early by testing CRC for MSI (adapted from Powell [112]).

5.2. Early identification of LS through screening polyps

Early identification of LS is highly desirable as the risk of developing CRC can be significantly reduced with increased cancer surveillance [123]. About 60% of CRC in LS cases are not diagnosed until after the age of 50 [124]. Thus, screening colorectal polyps obtained during colonoscopy that begins at 50 years of age could help identify LS patients and at-risk family members before cancer develops.

Screening for MSI in colon polyps could shift LS diagnosis earlier, allowing for earlier monitoring and improved chances of preventing cancer. However, colorectal polyps exhibit a milder MSI phenotype compared to more advanced neoplasms, limiting adoption of this strategy. Estimates for the incidence of MSI in LS adenomas range from 41 to 86% (average of 70%), which is comparable to IHC sensitivity of 49–82% (average of 72%) [125–131]. A study by Yurgelun and colleagues [130] found that while the overall MSI detection rate in adenomatous polyps from individuals with known pathogenic MMR mutations was 54%, all polyps larger than 10 mm in size exhibited MSI-H and loss of MMR expression by IHC. The higher level of MSI in the larger polyps is likely due to stepwise nature of MSI, in which larger deletions result from multiple smaller sequential replication errors that accumulate through-
out many cell divisions [132]. This phenomenon might explain why it is more difficult to detect MSI in small polyps as they would undergo fewer cell divisions after loss of MMR activity. Despite this, MSI can occur at a very early stage of adenoma formation, as it has been found in aberrant crypt foci of microscopic size [133, 134] and has even been observed in normal colonic mucosa of patients with LS [135].

Figure 5. Mutation frequency in mononucleotides increases exponentially with repeat length. Markers with poly-A tracts of 20–30 base pairs are currently used for MSI testing. Markers over 40 base pairs exhibit higher mutability and, therefore, increased MSI sensitivity.

Increasing the sensitivity of MSI testing could facilitate screening adenomas for early identification of LS patients. Bacher and colleagues compared the sensitivity of microsatellite markers with very long poly-A runs of 40–60 base pairs with currently used markers for MSI testing. The long mononucleotide repeat markers were identified from BLAST searches of human genome databases and the frequencies of insertion/deletion mutations were compared to existing markers with shorter poly-A tracks [136, 137]. Mutation frequencies were found to increase exponentially with increasing repeat length (Figure 5) in agreement with other studies of microsatellites [13, 15, 138, 139]. This finding is significant as mutation frequencies can serve as a surrogate for MSI sensitivity.

To determine whether the detection of MSI in colorectal polyps could be increased using long mononucleotide repeat markers, 430 polyps from 160 patients were screened using the Bethesda panel, MSI Analysis System (Promega Corporation, Madison, United States), and an experimental panel of long mononucleotide repeats (Promega Corporation, Madison, United States) (Figure 6) [140]. Using the long mononucleotide repeat panel, 15 tumors were scored as MSI-H compared to nine for the Bethesda panel and eight for the MSI Analysis System. This difference represented a 1.7–1.9-fold increase in relative sensitivity for the detection of MSI-H polyps over currently used markers. Importantly, a high proportion (80%) of MSI-H polyps was likely from LS patients. The relative MSI sensitivity of the long mononucleotide repeat markers was higher than any markers in the Bethesda panel and the MSI Analysis System (Figure 7). The sensitivity and specificity for the detection of MMR-deficient lesions were estimated based on IHC data on MMR protein expression (Table 4). The sensitivity and
specificity were 100 and 96% for the long mononucleotide repeat panel compared to 67 and 100% for the MSI Analysis System and 75 and 97% for the Bethesda panel. The difference in sensitivity between the long mononucleotide repeat panel and the other panels was statistically significant.

Figure 6. Prototype MSI Analysis System 2.0 (Promega Corporation, Madison, United States). The multiplex contains both short mononucleotide repeat markers, NR-21, BAT-25, BAT-26, and MONO-27, and new long mononucleotide repeat markers BAT-52, BAT-56, BAT-59, BAT-60, and pentanucleotide repeats Penta-C and Penta-D. The short mononucleotide repeats have a proven track record for MSI testing of CRC and the longer mononucleotide repeats have increased sensitivity for detection of MSI in colon polyps and extra-colonic tumors that often exhibit attenuated phenotypes. The pentanucleotide repeats are included to confirm sample identity.

Figure 7. The relative MSI sensitivity of the long mononucleotide repeat (LMR) markers. The percentage of MSI-H tumors that were MSI-positive for each individual marker was determined. The sensitivity of the LMR panel was significantly higher than the Bethesda panel ($p < 0.0038$) and the MSI Analysis System ($p = 0.0012$) using the $t$-test.
Table 4. Sensitivity and specificity of new long mononucleotide repeats (LMR).

<table>
<thead>
<tr>
<th>Marker</th>
<th>True positive</th>
<th>False negative</th>
<th>True negative</th>
<th>False positive</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMR panel</td>
<td>12</td>
<td>0</td>
<td>67</td>
<td>3</td>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>MSI analysis system</td>
<td>8</td>
<td>4</td>
<td>75</td>
<td>0</td>
<td>67</td>
<td>100</td>
</tr>
<tr>
<td>Bethesda panel</td>
<td>9</td>
<td>3</td>
<td>66</td>
<td>2</td>
<td>75</td>
<td>97</td>
</tr>
</tbody>
</table>

The use of the long mononucleotide repeat markers increased confidence in the MSI scoring as a consequence of a higher number of MSI-positive markers and larger allelic size changes for a given sample. MSI analysis with the long mononucleotide repeat panel resulted in MSI-H samples typically (80% of cases) exhibiting instability in four out of five or five out of five markers. With one exception, these cases also exhibited loss of MMR expression by IHC, had a germline MMR mutation, or both. Moreover, the significantly larger size changes in long mononucleotide repeats further simplified MSI classification by reducing the number of ambiguous calls often associated with small changes in the allele size that are observed when assaying shorter mononucleotide repeat sequences (Figure 8). The results of this study indicate that these new long mononucleotide repeat markers can increase sensitivity for the detection of MSI in polyps to a level approaching that reported in the literature for CRC with current marker systems. This increased sensitivity opens the possibility of screening polyps for an early detection of LS, while further study will be needed to be fully confident in these results and conclusions.

Figure 8. Long mononucleotide repeat markers for MSI analysis typically result in larger allele size changes in MSI-H tumors. Electropherogram of MSH2-deficient colon tumor and matching normal tissue screened for MSI using the experimental long mononucleotide repeat panel which shows all five markers were unstable with size shifts of up to 23 base pairs.
5.3. Alternative methods for LS testing

Current PCR-based MSI testing utilizes a small, standardized panel of highly unstable mononucleotide repeat markers to detect loss of MMR function. An alternative approach for MSI testing is to use highly scalable next generation DNA sequencing (NGS) technologies to infer MSI status. The main advantages of NGS are that multiple targets can be tested simultaneously, more efficiently, more cost effectively, and with higher sensitivity than with traditional Sanger sequencing. The main disadvantages are the greatly increased complexity of results and the return of uncertain or unexpected findings. Because the majority of MSI-positive tumors are due to epigenetic changes rather than genetic changes in MMR genes, even sequencing all MMR genes by NGS will not reliably infer MSI status in a tumor. To address this limitation, Hempelmann and colleagues [141] used NGS to sequence the five standard mononucleotide repeat loci in the MSI Analysis Kit (Promega Corporation, Madison, United States) to determine tumor MSI status. Using NGS they analyzed 81 CRC specimens (44 MSI-H and 37 MSI stable) previously subjected to PCR-based MSI testing. The MSI status of 95% of the specimens was interpretable by NGS and all but four samples were concordant with previous MSI classification. The samples generating ambiguous results were repeated and the result was the same, indicating that the NGS assay may not confidently infer MSI status for a small fraction of samples. While the NGS approach did not substantially improve sensitivity or specificity over existing assays, NGS offers an advantage of automated analysis based on quantitative, descriptive statistics which the authors suggest may improve intra- and interlaboratory variation.

Another approach to diagnose LS is direct sequencing of the MMR genes without previous screening with MSI or IHC. This approach simplifies the traditional multi-step testing procedure, but greatly increases the number of cases receiving costly germline MMR sequencing. Moreover, germline mutations in MMR genes may not be found in up to 30% of suspected LS cases [43]. Heritable, constitutional epimutations in MLH and MSH2 explain many of these cases [142, 143]. Biallelic somatic mutations in MMR genes may account for up to 60–70% of germline MMR-negative cases [144]. Another potential limitation of direct sequencing of MMR genes is the high number of variants of unknown clinical significance, which account for around one third of germline MMR mutations [145]. The International Society of Gastrointestinal Hereditary Tumors currently reports a total of 3104 MMR gene variants (1198 for MLH1, 1098 for MSH2, 547 for MSH6, and 261 for PMS2) [146]. Communicating test results showing variants of unknown significance to patients can be challenging due to the potential psychological impact of reporting uncertain test results. Since both LS and MSI are caused by MMR defects, screening with MSI serves as a surrogate marker of LS and is a functional test for loss of MMR. Determining tumor MSI status also provides a prognostic and therapeutic value for individualizing treatment not only for LS patients, but also for those with sporadic MSI-H CRC lacking a germline MMR mutation [5, 6].

5.4. Distinguishing Lynch syndrome from non-Lynch syndrome CRC

There are multiple types of non-Lynch syndrome CRC that can mimic the disease and confound diagnosis [144, 147]. Many of these tumors are MSI-positive or show loss of MMR
gene expression by IHC, but lack germline mutations [144, 148]. Distinguishing these mimics from LS is clinically important, as treatment and surveillance for these patients and their at-risk family members differ.

Nonfamilial LS mimics include sporadic MSI-positive CRC and Lynch-like syndrome (LLS) cancers. Hyper-methylation of the MLH1 gene is responsible for about 80% of cases where MLH1 is missing without MLH1 germline mutations. These sporadic MSI-positive CRC are fairly easy to distinguish from LS because of older age of onset, lack of family history of cancer, the presence of BRAF V600E mutation, and/or methylation of MLH1. More challenging are cases where the LLS cancers exhibit MSI and loss of MMR expression, but patients lack germline MMR mutations [149]. Mutations in EPCAM explain about 20–25% of LLS cases which show loss of MSH2 expression but no germline MSH2 mutation. Deletions in the EPCAM gene lead to hypermethylation of the MSH2 promoter and subsequent MSH2 silencing. Most (70%) of the remaining unexplained LLS cases have cancers with biallelic somatic MMR mutations [148]. Thus, the distinguishing features of LLS are an MSI-positive phenotype and somatic biallelic MMR gene mutations. LLS has also been shown to occur in some endometrial cancers [148].

Familial CRC mimics include (1) polymerase proofreading associated polyposis (PPAP) caused by mutations in POLE or POLD1, (2) familial colorectal cancer type X (FCCTX) of unknown etiology, (3) germline MLH1 methylation, and (4) constitutional mismatch repair-deficiency (CMMRD) caused by biallelic germline MMR mutations. PPAP is a rare inherited form of CRC that is caused by germline mutations in POLE (encoding DNA polymerase ε) or POLD1 (encoding DNA polymerase δ) [150, 151]. Individuals with PPAP can develop CRC as early as 20 years of age and POLD1 mutation carriers are also at increased risk for endometrial and brain cancers. Tumors from PPAP individuals are MSI stable even though they have 100-fold more mutations in nonrepetitive DNA than sporadic MSI-positive tumors [23]. The absence of MSI in these CRCs is a distinguishing feature of PPAP. Another type of familial CRC lacking MSI and germline MMR mutations is familial colorectal cancer type X (FCCTX) [152]. The genetic cause of FCCTX is unknown. FCCTX individuals have about twofold increased risk of CRC compared to the general population, but do not develop other LS-spectrum cancers. Methylation of MLH1 is usually associated with sporadic MSI-positive CRC and is not heritable. However, in rare cases inherited germline epigenetic silencing of MLH1 has been reported to predispose to cancer development in a pattern typically found in LS families [153]. CRC and other tumors from individuals with MLH1 germline epimutations exhibit MSI and lack of MLH1 expression. Diagnosis of MLH1 germline epimutations is accomplished by methylation analysis of tumor and germline samples. Constitutional mismatch repair-deficiency (CMMRD) is another rare disorder that is caused by biallelic germline mutations in MMR genes (most commonly PMS2 and MSH6) that predisposes them to childhood cancers [154]. CMMRD individuals may present with CRC, brain tumors and/or leukemia and lymphoma. These tumors exhibit MSI and loss of MMR protein expression like LS, but can be distinguished by presence of biallelic germline MMR gene mutations. Screening CRC tumors for MSI followed by germline MMR sequencing is an effective strategy to distinguish LS from these non-LS diseases.
5.5. Use of MSI as a predictive biomarker

MSI-positive CRC is associated with a better prognosis and a decreased likelihood of metastasis to lymph nodes and distant organs [155]. A meta-analysis with 7642 cases clearly demonstrated that patients with MSI-H tumors have a significantly better prognosis than those with MSS tumors (hazard ratio for death = 0.65) [156]. There is growing evidence that the improved prognosis of MSI-positive tumors is due to the accumulation of frame shift mutations in genes containing coding microsatellites [157]. Translation of proteins with mutation-induced frame shift peptides renders MSI cancers highly immunogenic, allowing the body’s immune system to more effectively target cancer cells.

Figure 9. MSI is a biomarker for PD-1 blockade immunotherapy. Increased mutational burden in MMR-deficient tumors creates mutation associated neo-antigens (MANA) responsible for the immune response. Tumors with MSI evade immune surveillance by upregulation of immune checkpoint proteins programmed cell death 1 (PD-1) and its ligand PD-L1. Interaction of the PD-1 co-receptor and its ligand PD-L1 suppresses the immune system, and antibodies to either molecule have shown promise for reinvigorating the immune system, allowing T-cells to target and destroy cancer cells. Ab, antibody; MHC, major histocompatibility complex; TCR, T cell receptor; *, mutation.

While MSI status is a good prognostic factor for CRC, its predictive value for chemosensitivity remains controversial. The initial study on the use of 5-FU-based adjuvant chemotherapy by Ribic and colleagues [158] found that patients with advanced stage MSI-negative CRC benefited, but patients with MSI-H CRC did not. A number of subsequent clinical studies have confirmed these results [159, 160]. The clinical results are supported by in vitro evidence showing that MutSα and/or MutSβ binds to 5-FU incorporated DNA resulting in cell death, indicating that a functioning MMR system is required for the cytotoxic effect of 5-FU [161]. In contrast, a number of studies have failed to find any effect of MSI status on 5-FU treatment response [162–164]. A recent meta-analysis involving 9212 patients concluded that there was
no clear difference in response to treatment based on MSI status [164]. However, the evidence for a detrimental effect of 5-FU treatment on MSI-positive tumors was sufficiently strong to justify another clinical trial (ClinicalTrials.gov identifier: NCT00217737; this study is ongoing) to assess the role of MSI in predicting response to adjuvant chemotherapy.

One of the most promising new approaches for treating advanced CRC is immune checkpoint therapy, which activates the body’s natural antitumor activity (Figure 9) [5, 6]. Immune checkpoint therapy is less toxic than chemotherapeutic regimens and has potential for durable responses in advanced cancer patients who may otherwise only live a few months. It is estimated that approximately 50% of CRC in patients will progress to metastatic cancer. Prognosis for advanced CRC remains poor with overall 5-year survival at 70% for patients with localized lymph node metastases and 13% for patients with organ metastases.

Immune surveillance can effectively recognize and eliminate cancerous cells and is regulated by a balance between stimulatory and inhibitory signals (i.e., immune checkpoints). Under normal conditions, immune checkpoints are inhibited to maintain self-tolerance and avoid inappropriate overreaction, such as an auto-immune disease. In the presence of tumor cells, immune surveillance is activated. Selection pressure exerted by the immune system on tumor cells can lead to resistant clones that survive by inhibiting immune surveillance. MSI-positive cancers exhibit active immune response due to high number of neo-antigens that are produced by frameshift mutations in coding repeats in MMR-deficient cells. High expression of checkpoint molecules in MSI CRC creates an immunosuppressive microenvironment that is thought to help MSI tumors evade immune destruction by the infiltrating immune cells. Clinical trials of stage IV CRC with anti-PD-1 antibody pembrolizumab have been shown to be promising for reinvigorating the immune system to target and destroy cancer cells (Figure 10) [5]. MSI was found to be a significant predictor of the progression-free survival rate of 78% for MMR deficient CRC, 67% for MMR-deficient non-CRC cancer, and 11% in MMR-proficient CRC.

![Figure 10](https://example.com/figure10.png)

**Figure 10.** Clinical benefit of anti PD-1 antibody pembrolizumab treatment according to MSI status. Kaplan-Meier curves are shown for progression-free survival in the cohorts with CRC (reprinted by permission from Le et al. [5]).
6. Summary and concluding remarks

The vast majority of the estimated one million individuals with LS in the United States are not diagnosed. Early identification of individuals with LS is critical as the risk of developing cancer can be significantly reduced with increased surveillance. It is now recognized that screening strategies which rely on clinical criteria alone for the diagnosis of LS lack the needed sensitivity and that new strategies are required to address the underdiagnoses of the disease. The medical and life costs related with missed diagnosis are substantial due to the high costs and poor prognosis associated with treating advanced cancers. In an effort to increase detection of LS, there has been a growing support for universal screening of all new colorectal and endometrial cancers. Since definitive diagnosis of LS requires expensive germline MMR mutation analysis, cost-effective strategies are needed to prescreen for possible LS patients to triage those who will need germline analysis. In 1993, MSI became the first biomarker to be used for the detection of LS. Subsequent improvements, such as the change to all mononucleotide repeats and the introduction of fluorescent multiplex PCR methodology, have made MSI a highly accurate and cost-effective biomarker for LS (Figure 11). New technologies for MSI detection, like next generation sequencing, open the possibility of a single test for LS that determines tumor MSI status and MMR germline mutations. MSI is currently an important prognostic and diagnostic biomarker for LS, but it is poised to take on a much greater role in prediction of responses to the new immunotherapies targeted at MSI-positive tumors.

Figure 11. Time line of important events in the development of MSI testing for LS [1, 6, 8, 17, 25, 37, 44, 94, 98, 100, 118, 122, 124, 141, 165–174].
Author details

Jeffery W. Bacher1,3, Linda Clipson2, Leta S. Steffen1 and Richard B. Halberg2,3,4*

*Address all correspondence to: rbhalberg@medicine.wisc.edu

1 Research and Development, Promega Corporation, Madison, WI, USA
2 McArdle Laboratory for Cancer Research, Department of Oncology, University of Wisconsin, Madison, WI, USA
3 Division of Gastroenterology & Hepatology, Department of Medicine, University of Wisconsin, WI, USA
4 Carbone Cancer Center, University of Wisconsin, Madison, WI, USA

References


[18] Loeb LA, Springgate CF, Battula N. Errors in DNA Replication as a Basis of Malignant Changes. Cancer Res. 1974;34:2311–2321


[22] Loeb LA. Mutator Phenotype May Be Required for Multistage Carcinogenesis. Cancer Res. 1991;51:3075–3079


[64] Zighelboim I, Goodfellow PJ, Gao F, Gibb RK, Powell MA, Rader JS, Mutch DG. Microsatellite Instability and Epigenetic Inactivation of MLH1 and Outcome of Patients


and Seven Other Polymerase Genes in 266 Patients with Colorectal Adenomas and Carcinomas. Int J Cancer. 2014;137:320–331. DOI: 10.1002/ijc.29396


and Survival in Patients with Microsatellite-Unstable Colorectal Cancer. Gastroenterol. 2004;126:394–401


