DNA mismatch repair proteins: scientific update and practical guide

Adrian C Bateman

ABSTRACT

DNA mismatch repair (MMR) proteins are essential for the recognition and correction of sporadic genetic mutations that occur during DNA replication. Deficient MMR function (dMMR) leads to an increased risk of development of neoplasia. Identification of dMMR within tumours can suggest a high chance of the inherited cancer condition Lynch syndrome and predicts poor clinical response to certain conventional chemotherapies but an increased likelihood of response to immunotherapy. This review provides an update on the biology of MMR proteins, their encoding genes and mechanisms for the development of dMMR. This is followed by a discussion of the identification and significance of dMMR in routine clinical practice.

MISMATCH REPAIR PROTEINS AND THEIR FUNCTION

Mismatch repair (MMR) proteins are essential for repairing DNA errors (eg, point mutations) that are generated during DNA replication. There are at least seven MMR proteins in humans, of which four have the most clinical relevance in human cancer biology—MLH-1, MSH-2, MSH-6 and PMS-2. Mismatch repair proteins are arranged as heterodimers in which MLH-1 associates with PMS-2 and MSH-2 associates with MSH-6. The MLH-1/PMS-2 pair recognises mismatched nucleotide base pairs and initiates repair, while the MSH-2/MSH-6 pair acts as an endonuclease. These proteins are encoded by their corresponding genes (MMR genes)—MLH-1, MSH-2, MSH-6 and PMS-2.

CONSEQUENCES OF LOSS OF MMR PROTEIN FUNCTION

Loss of function of one or more MMR proteins (deficient MMR (dMMR)) leads to impaired DNA repair capability. This results in the accumulation of spontaneous genetic mutations across the genome, affecting the function of many genes. Those involved in normal cellular growth and differentiation (proto-oncogenes) may become amplified (present at multiple copy numbers) or undergo gain-of-function mutations (creating oncopgenes), while the function of tumour suppressor genes may be impaired. This leads to an increased risk of development of neoplasia. Identification of dMMR has two main areas of clinical relevance—screening for inherited cancer syndromes, for example, Lynch syndrome and the prediction of response to conventional chemotherapy and immunotherapy. Lynch syndrome was first described by the American physician Henry T Lynch in 1966 and is a familial cancer condition that is almost always associated with the presence of a mono-allelic germline mutation in an MMR gene. Affected patients are at increased risk of development of one or more of a range of neoplasms (table 1). The syndrome was originally termed ‘hereditary non-polyposis colorectal cancer (CRC)’ due to its common association with CRC. However, recognition of the broad range of possible tumour types has resulted in its renaming as ‘Lynch syndrome’. Around 50% of patients with Lynch syndrome have neoplasms other than CRC or endometrial carcinoma. Certain other tumours may rarely be associated with Lynch syndrome but the evidence supporting their inclusion within the spectrum of Lynch-associated neoplasms is currently weak, for example, breast cancer.

MECHANISMS OF DMMR DEVELOPMENT

dMMR function can occur due to a mutation within an MMR gene or because of inactivation of an otherwise intact MMR gene. MMR gene mutations may be present in germline DNA or can occur as somatic events within a tumour. Mono-allelic germline MMR gene mutations are associated with Lynch syndrome but bi-allelic germline mutations may very rarely be encountered and lead to ‘constitutional MMR deficiency’ or ‘childhood cancer syndrome’.

Inactivation of MMR genes occurs with MLH-1 and is a consequence of hypermethylation of the promoter sequence of the gene. Hypermethylation is a common biological process that comprises the addition of methyl groups (–CH3) to CpG islands. The latter are DNA regions where a cytosine nucleotide is followed by a guanine nucleotide at a high frequency—the ‘p’ indicates that the two nucleotides are separated by only a phosphate group that is, they are present on a single strand of DNA. Hypermethylation is a physiological process for controlling gene expression and many sites within the genome are methylated during cellular development and differentiation. However, most CpG islands within promoter sequence DNA are usually unmethylated. Hypermethylation of promoter sequences causes inactivation of the promoter and switches off the corresponding gene. Tumours such as CRC and precursor lesions showing these genetic changes are described as ‘CIMP-high’ (ie, CpG island methylation phenotype—high level).

Hypermethylation is usually a somatic process and within colorectal neoplasia, occurs commonly in the serrated pathway of carcinogenesis in association with BRAF mutations. Constitutional
IDENTIFICATION OF DMMR IN DIAGNOSTIC PRACTICE

In clinical practice, dMMR may be detected at a genetic, protein or functional level.Diagnostic histopathologists will be most familiar with the use of immunohistochemistry (IHC) to demonstrate the presence of MMR proteins within tumour cell nuclei. Due to MMR protein heterodimerism, the pattern of loss of expression of the four most commonly studied proteins provides useful clues to the most likely underlying genetic defect (table 2).

Table 2 Patterns of MMR protein loss identified on IHC

<table>
<thead>
<tr>
<th>Pattern of expression loss on IHC</th>
<th>Likely clinical implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common patterns</td>
<td></td>
</tr>
<tr>
<td>MLH-1 and PMS-2</td>
<td>MLH-1 inactivation</td>
</tr>
<tr>
<td></td>
<td>MLH-1 mutation</td>
</tr>
<tr>
<td>MSH2 and MSH-6</td>
<td>MSH-2 mutation</td>
</tr>
<tr>
<td></td>
<td>EPCAM mutation (less common)</td>
</tr>
<tr>
<td>MSH-6 alone*</td>
<td>MSH-6 mutation</td>
</tr>
<tr>
<td></td>
<td>MSH-2 mutation (less common)</td>
</tr>
<tr>
<td>PMS-2 alone†</td>
<td>PMS-2 mutation</td>
</tr>
<tr>
<td></td>
<td>MLH-1 mutation (less common)</td>
</tr>
<tr>
<td>Rare patterns</td>
<td></td>
</tr>
<tr>
<td>MLH-1 alone</td>
<td>MLH-1 inactivation</td>
</tr>
<tr>
<td></td>
<td>MLH-1 mutation</td>
</tr>
<tr>
<td>MSH-2 alone</td>
<td>MSH-2 mutation</td>
</tr>
<tr>
<td></td>
<td>EPCAM mutation</td>
</tr>
</tbody>
</table>

*MSH-2 can form heterodimers with MMR proteins other than MSH-6 and therefore MSH-2 expression is usually not lost when an MSH-6 mutation is present.
†MLH-1 can form heterodimers with MMR proteins other than PMS-2 and therefore MLH-1 expression is usually not lost when a PMS-2 mutation is present.

It has been suggested that an IHC panel limited to MSH-6 and PMS-2 may be a more cost-effective way of detecting dMMR.3–10 As well as identifying patients with a somatic or germline MSH-6 or PMS-2 mutation, loss of PMS-2 or MSH-6 expression also occurs in those with a somatic or germline MLH-1 or MSH-2 mutation (or MLH-1 inactivation), respectively. Isolated loss of MLH-1 or MSH-2 expression is rare. If this two-antibody panel is used, IHC for the ‘paired’ MMR protein could be performed if loss of MSH-6 or PMS-2 is found, in order to determine which MMR gene is most likely to be defective. In practice, many laboratories still include all four of these MMR proteins in their IHC panel as the overall cost is still relatively low and the ability to assess both MMR proteins within each ‘pair’ can be useful when the staining quality is suboptimal, for example, due to poor initial tissue fixation.

Some of the more commonly encountered difficulties in the interpretation of MMR protein expression are described in table 3. In Lynch (and ‘Lynch-like’) syndrome, loss of MMR protein expression usually occurs in a widespread and homogeneous manner within the tumour. However, heterogeneity of MMR protein expression may be encountered and manifests as a defined area of lost expression in the presence of retained stromal and inflammatory cell MMR protein staining—within a tumour that otherwise shows widespread expression of the protein(s) concerned. In this situation, an individual malignant gland can sometimes show zones of retained and lost MMR protein expression. Assuming that technical reasons for variations in staining intensity have been excluded, this heterogeneous staining implies the presence of a clonal somatic MMR gene mutation or localised hypermethylation process, that is, a mutation or genetic inactivation affecting only a small area of the tumour. Importantly, it does not imply that a germline mutation is present.11

It has been suggested that an IHC panel limited to MSH-6 and PMS-2 may be a more cost-effective way of detecting dMMR.3–10 As well as identifying patients with a somatic or germline MSH-6 or PMS-2 mutation, loss of PMS-2 or MSH-6 expression also occurs in those with a somatic or germline MLH-1 or MSH-2 mutation (or MLH-1 inactivation), respectively. Isolated loss of MLH-1 or MSH-2 expression is rare. If this two-antibody panel is used, IHC for the ‘paired’ MMR protein could be performed if loss of MSH-6 or PMS-2 is found, in order to determine which MMR gene is most likely to be defective. In practice, many laboratories still include all four of these MMR proteins in their IHC panel as the overall cost is still relatively low and the ability to assess both MMR proteins within each ‘pair’ can be useful when the staining quality is suboptimal, for example, due to poor initial tissue fixation.

Some of the more commonly encountered difficulties in the interpretation of MMR protein expression are described in table 3. In Lynch (and ‘Lynch-like’) syndrome, loss of MMR protein expression usually occurs in a widespread and homogeneous manner within the tumour. However, heterogeneity of MMR protein expression may be encountered and manifests as a defined area of lost expression in the presence of retained stromal and inflammatory cell MMR protein staining—within a tumour that otherwise shows widespread expression of the protein(s) concerned. In this situation, an individual malignant gland can sometimes show zones of retained and lost MMR protein expression. Assuming that technical reasons for variations in staining intensity have been excluded, this heterogeneous staining implies the presence of a clonal somatic MMR gene mutation or localised hypermethylation process, that is, a mutation or genetic inactivation affecting only a small area of the tumour. Importantly, it does not imply that a germline mutation is present.11

Msats are non-coding DNA regions that are present throughout the genome and that—along with coding regions—are replicated imperfectly during cell division if there is dMMR. These imperfections can be identified within tumour-derived DNA using PCR-based testing and this phenomenon is termed ‘microsatellite instability’ (MSI). MSI therefore represents an
alternative means of detecting dMMR but the technique does not indicate which MMR gene is most likely to be involved. MMR protein IHC and PCR-based MSI testing are usually considered as interchangeable methods for screening tumours for dMMR. However, if the index of suspicion for a germline MMR gene defect is high, for example, in a patient under investigation by a clinical genetics department for possible Lynch syndrome, they can both be applied to the same tumour. This approach can minimise the risk of missing dMMR—this could otherwise occur for example, due to the presence of a mutation that affects the function of an MMR protein but in which the enzyme is still demonstrable using IHC, or the presence of a defect in a different, rarely implicated MMR gene, for example.

Next generation sequencing is becoming more commonly used as an alternative method for detecting MSI. This is often as part of a broader panel for identifying somatic mutations and can be performed on formalin-fixed and paraffin-embedded tissue-derived DNA. This method may be better suited to large scale analyses and does not require the interrogation of matched normal tissue alongside the tumour sample.12

Formal genetic mutation analysis of germline DNA is the gold standard for identifying MMR gene mutations and if found, represent a powerful tool for screening family members. However, sometimes rare or novel mutations are found and these can require correlation with IHC and MSI testing in order to assess whether they are likely to be pathogenic (ie, disease-causing) in nature.

A germline MMR gene mutation confirming the diagnosis of Lynch syndrome is found in almost all patients with an appropriate family history and tumour characteristics suggesting the presence of such a mutation. Constitutional MLH-1 promoter hypermethylation may account for up to 3% of cases of Lynch syndrome, especially in patients where there is no clear family history.13 The presence of dMMR when no germline MMR gene defect is found despite the presence of supporting clinical and tumour features has been termed ‘Lynch-like syndrome’.14 The mechanisms underlying dMMR in this situation are varied and could include the presence of (1) an undetectable MMR gene mutation, (2) a mutation in non-MMR gene that can cause MSI or (3) a somatic bi-allelic MMR gene mutation or hypermethylation.13

**EPCAM MUTATIONS AND DMMR**

In some patients with Lynch syndrome, IHC demonstrates loss of MSH-2 expression within their tumours but a germline mutation in MSH-2 is not identified. Instead, a germline mutation is present at the 3’ end of the *EPCAM* gene (the gene encoding epithelial cell adhesion molecule - EpCAM), which results in hypermethylation of the MSH-2 promoter sequence and inactivation of MSH-2.16 This mechanism accounts for 20%–25% of patients whose CRC show loss of MSH-2 expression but in whom no germline MSH-2 mutation is detected—or 2%–3% of all Lynch syndrome cases.17 Bi-allelic *EPCAM* mutations are associated with loss of *EPCAM* expression on IHC in Lynch syndrome-associated tumours.18

**INVESTIGATION OF LOSS OF MLH-1 EXPRESSION IN CLINICAL PRACTICE**

Loss of MLH-1 expression—usually accompanied by loss of PMS-2 expression—within tumours is the most commonly encountered abnormality on MMR protein IHC. This can indicate the presence of a germline MLH-1 mutation but much more commonly occurs due to somatic MLH-1 inactivation associated with hypermethylation of its promoter sequence. In CRC, DNA from tumours showing MLH-1 loss is subjected to *BRAF* mutation analysis. If a *BRAF* mutation is found, the tumour is most likely to be sporadic in nature and has probably developed along the serrated pathway. If no *BRAF* mutation is found, tumour DNA then undergoes MLH-1 promoter hypermethylation analysis. The presence of hypermethylation almost always indicates that the tumour is sporadic. The exception to this is constitutional hypermethylation that is, hypermethylation of germline DNA, representing a very rare cause of Lynch syndrome. In contrast, endometrial cancers lacking MLH-1 expression rarely show *BRAF* mutations and therefore if MLH-1 loss is found in these tumours, direct progression to MLH-1 promoter hypermethylation analysis is required. In the context of absent MLH-1 and PMS-2 expression and in CRC if no *BRAF* mutation is present, patients with CRC or endometrial cancer in which no MLH-1 promoter hypermethylation is found should be referred for genetic testing. This is because in these settings, they are likely to possess a germline MLH-1 mutation.

**MMR PROTEIN IHC IN BIOPSIES VERSUS RESECTIONS IN CRC**

In CRC, histopathologists are usually encouraged to perform MMR protein IHC on biopsy material, if possible. This can provide an MMR status earlier in the patient pathway and rapid formalin fixation of small biopsy tissue fragments results in optimised antigen preservation and therefore minimised technical artefacts on IHC. However, sometimes IHC on biopsies is not possible for example, if the biopsy shows features that are not diagnostic of adenocarcinoma, or if an emergency colorectal resection is performed without prior biopsy for example, due to bowel perforation or obstruction. In these situations, testing must instead be undertaken on resection material.

Caution must be applied when interpreting MMR protein IHC in resection specimens from patients who have undergone neo-adjuvant therapy. The latter can lead to loss of MMR protein expression in the absence of an associated germline MMR mutation and is seen most commonly with MSH-6. Therefore, if loss of expression is found when examining a post-neoadjuvant therapy specimen, repeat MMR IHC on pretreatment biopsy material is required if this is available.19

**MMR PROTEIN IHC IN COLORECTAL ADENOMAS VERSUS CRC**

In colorectal neoplasia, while IHC and MSI testing for dMMR is usually undertaken on biopsy or resection material showing adenocarcinoma, useful information may sometimes be gained in patients under investigation for possible Lynch syndrome, via IHC performed on colorectal adenomas. This testing is sometimes requested by clinical genetics departments, for individuals from possible Lynch syndrome families who have had colorectal adenomas removed but who do not have CRC. In patients with known Lynch syndrome, colorectal adenomas show loss of MMR protein expression in around 80% of cases.20 Therefore IHC performed on adenomas can be informative if loss of MMR protein expression is demonstrated, but this test does not exclude the possibility of Lynch syndrome if expression is retained.

Loss of MMR protein expression can be seen in morphologically normal colonic mucosa in patients with Lynch syndrome but is only very exceptionally seen in individuals without this condition. This phenomenon can provide useful supporting evidence for the presence of Lynch syndrome, particularly in patients where germline MMR gene mutation analysis is inconclusive.21 22
NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE GUIDELINES FOR LYNCH SYNDROME SCREENING
The National Institute for Health and Care Excellence (NICE) published guidelines in 2017, indicating that all patients newly diagnosed with CRC should undergo screening for Lynch syndrome, with either MMR protein IHC or MSI testing as the initial investigative method. Until this time, MMR protein IHC was usually performed on CRC if the clinical features (eg, patient age <50 years) and/or the histological features of the tumour (eg, right-sided, poorly differentiated, mucinous or tumour-infiltrating lymphocyte-rich) raised the possibility of Lynch syndrome. Additional funding was not provided for pathology laboratories to support this development. Individual patient consent is not required for this screening process, unless and until direct testing for a germline mutation is needed. Implementation of these guidelines also requires a well-defined pathway for the further investigation of MLH-1-deficient CRC and a reliable mechanism for ensuring that patients in whom screening suggests the presence of Lynch syndrome are appropriately referred to clinical genetics departments. In 2020, NICE introduced similar guidelines for patients with endometrial cancer.

DMMR AND ONCOLOGICAL CANCER TREATMENT
In CRC, the presence of dMMR is associated with a better stage-adjusted prognosis compared to MMR-proficient tumours, but a reduced likelihood of response to certain conventional chemotherapy regimens, for example, 5-fluorouracil (5-FU). The reason for this lack of benefit from 5-FU therapy is unknown. However, patients with CRC showing dMMR (especially stage 2, ie, pT3N0 or pT4N0) should not receive adjuvant therapy with this agent. In contrast, patients with endometrial cancer showing dMMR may show an improved response to adjuvant radiotherapy.

The introduction of immunotherapy has improved clinical outlook within a range of cancers and the number of cancer types in which immunotherapy has been shown to be of benefit is ever-increasing. These treatments are based on boosting an antitumour immune response by patients’ own immune systems, usually by blocking molecular mechanisms that tumours use to evade host attack. The presence of dMMR leads to an increased mutational burden and the generation of novel peptide sequences by cancer cells, representing an enhanced range of epitopes that evade host attack. The presence of dMMR leads to an increased mutational burden and the generation of novel peptide sequences by cancer cells, representing an enhanced range of epitopes that are potentially recognisable by the host immune system. Therefore, tumours with dMMR may respond more favourably to immunotherapy than those lacking this feature.

Immunotherapies target cell surface molecules such as CTLA-4 and the PD-1/PD-L1 system. Of these, ipilimumab therapy for malignant melanoma, pembrolizumab therapy for non-small cell lung cancer and atezolizumab therapy for bladder cancer are already established treatments. More recent studies have demonstrated benefits of immunotherapy in a range of other advanced cancers with dMMR, for example, pembrolizumab in CRC (KEYNOTE-177 phase III trial) and in a range of non-CRCs including endometrial, gastric, cholangiocarcinoma and pancreatic (KEYNOTE-158 phase II trial).

CONCLUSION
MMR proteins play a critical role in DNA repair and therefore help protect against the accrual of sporadic mutations and the development of neoplasia. Lynch syndrome was originally described over 50 years ago but the factors leading to dMMR and its association with Lynch syndrome and ‘Lynch-like’ syndromes have been the subject of ongoing and intense study. Nowadays, a detailed understanding of the biology of dMMR is essential not only for the identification of inherited predisposition to cancer but also for guiding oncologists with treatment choices associated with both conventional chemotherapy and novel immunotherapies.

Handling editor Des Richardson.
Contributors I am the sole author of this script.
Funding The author has not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.
Competing interests None declared.
Patient consent for publication Not required.
Provenance and peer review Not commissioned; externally peer reviewed.

ORCID ID
Adrian C Bateman http://orcid.org/0000-0003-2222-4104

REFERENCES
Molecules in pathogenesis


