DNA mismatch repair, microsatellite instability and cancer

Minal Vaish. & B Mittal *

Department of Medical Genetics, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Raebareli Road, Lucknow 226 014, India

Mismatch (MMR) repair system plays a significant role in restoration of stability in the genome. Mutations in mismatch repair genes hamper their activity thus bring about a defect in mismatch repair (MMR) mechanism thereby conferring instability in the microsatellite sequences of both the coding and non-coding regions of the genome. Mutated mismatch repair genes result in the expansion or contraction of microsatellite sequence and confer microsatellite unstable or replication error positive phenotype. Hypermethylation of promoter regions of some of the MMR genes also causes inactivation of these genes and thus contribute to MSI. Microsatellite instability is an indicator of MMR deficiency and is a prime cause of varied tumorogenesis.

Genetic perturbation has been considered to be a hallmark of cancer. Genetic instability is manifested at the chromosome level as the incidence of aneuploidy, deletions, translocations and sister chromatid exchanges, whereas at the DNA level as altered DNA repair properties like gene amplification, deletion and point mutations. Genetic instability is a source of variability and is regarded as a driving force in the generation of variants with increased invasive and metastatic potential.

Tumors displaying high level of instability are a reflection of malfunction of DNA mismatch repair machinery. It is one of the multiple replication, recombination and repair processes that maintains genome integrity and is highly conserved among prokaryotes and eukaryotes. Mismatch repair system of the cell provides a protective role but mutations in mismatch repair (MMR) genes predispose a cell to tumor development.

Mismatch repair genes

Mismatch repair process is common to both prokaryotic and eukaryotic cells. It involves the recognition and resolution of incorrectly paired nucleotides in DNA. Its main role is in the rapid repair of replicative errors and provide the genome with 100-1000 fold level of protection against mutation and also guard genome by preventing recombination between non-homologous regions of DNA, thus maintains the genomic stability.

Mismatch repair genes were first described in bacteria and best characterized in Escherichia coli. Subsequently, they were identified in Saccharomyces cerevisiae and higher eukaryotes. The proteins that are involved in mismatch repair are of two types. The genes in first category are mut S, mut L and mut H that largely contribute in correcting replication errors. In higher eukaryotes mut H is not known (Table 1). The second category comprises proteins that participate in DNA metabolic pathways and these are helicases, endonucleases, polymerases and ligases.

Mut S is an ATPase and acts as homodimer. The locus of human homologue for mut S has been identified on chromosome 2p3. Six Mut S homologues (MSH1 to MSH6) have been identified in eukaryotes. MSH1 is required for normal mitochondrial function in S. cerevisiae and protects against base substitution. However, mammalian MSH1 homologue is not known to date. MSH4 and MSH5 have specialized roles in meiotic recombination in yeast and mice. MSH2, MSH3 and MSH6 [originally named as G: T binding protein (GTBP)] form two heterodimers-Mut Sα (MSH2/MSH6) and Mut Sβ (MSH2/MSH3) that recognize mispairs in DNA. Homologues of Mut S in higher eukaryotes (Xenopus laevis, mouse and human) are different from Mut S in the binding to the stretches of up to 14 extrahelical nucleotides (insertion deletion loops: IDLs) in vitro. Mut S and its homologue interact with DNA via their N-termini, while the C-terminal domains house dimerization and ATP binding domains. In vitro assays show that Mut S binds with high affinity to substrates containing base-base mispairs and insertion deletion loops (IDLs) that escape proofreading by the replicating polymerase.

*Correspondent author
TEL: 440004-8, ext. 2322; FAX: 91-0522-440 017
E-mail: balraj@sgpgi.ac.in; bml_pgi@yahoo.com
The homologues of both Mut H and Mut S show homology to bacterial damage specific glycosylases/lyases, it is involved in base excision repair.

Table I — Mismatch repair genes in prokaryotes and eukaryotes

<table>
<thead>
<tr>
<th>Prokaryote E. coli</th>
<th>Eukaryote Yeast</th>
<th>Human</th>
<th>Chromosomal location in human</th>
</tr>
</thead>
<tbody>
<tr>
<td>mut S</td>
<td>MSH1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>mut S</td>
<td>MSH2</td>
<td>MSH2</td>
<td>2p</td>
</tr>
<tr>
<td>mut S</td>
<td>MSH3</td>
<td>MSH3</td>
<td>5q</td>
</tr>
<tr>
<td>mut S</td>
<td>MSH4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>mut S</td>
<td>MSH5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>mut L</td>
<td>—</td>
<td>MSH6/GTBP</td>
<td>2p</td>
</tr>
<tr>
<td>mut L</td>
<td>MLH1</td>
<td>MLH1</td>
<td>3p</td>
</tr>
<tr>
<td>mut L</td>
<td>MLH2, MLH3</td>
<td>PMS1</td>
<td>7q</td>
</tr>
<tr>
<td>mut L</td>
<td>—</td>
<td>PMS2</td>
<td>2q</td>
</tr>
<tr>
<td>mut L</td>
<td>—</td>
<td>PMS1</td>
<td>—</td>
</tr>
<tr>
<td>mut U</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Positional cloning strategies identify human mut L homologue on chromosome 3p9. Human homologue of post meiotic segregation-hPMS2 has been identified on 7q whereas PMS1, which is S. cerevisiae homologue of hPMS2, is also known10. These factors exist as heterodimer of either hMLH1 and hPMS2 or MLH1 and PMS111. The precise biochemical role of eukaryotic Mut L homologues in mismatch repair is unclear. The role of S. cerevisiae MLH1-PMS1 complex is identified as to enhance the binding of MSH2/MSH3 heterodimer to an insertion deletion loop substrate and in humans Mut L homologues interact with Mut S homologues on a DNA substrate and show ATP dependent activity12.

Endonuclease Mut H has a role to discriminate between the templates and newly replicated strands and introduces a nick in the nascent strand at hemimethylated GATC sequences in the prokaryote13. Mut U plays a key role in the repair process. Mut L-Mut U interaction may control the directionality of the unwinding process and proceed towards the mispair14. The homologues of both Mut H and Mut U are not known in eukaryotic system.

DNA mismatch repair process

Mismatch recognition is accomplished by the binding of Mut S, a dimer to DNA substrates containing mismatches and its binding to DNA is followed by the binding of Mut L which is also a dimer and increases the stability of Mut S-DNA mismatch complex5,16. It results in the formation of symmetrical loop which increases in size until reaches hemimethylated site where Mut H, a single stranded endonuclease recognizes and nicks unmethylated or nascent strand at the hemimethylated (d GATC) sites17. This is followed by unwinding of DNA strand containing mismatch by Uvr D (Mut U)14. Mut L loads these Uvr D molecules.

The loop remains intact and Mut S-Mut L binds to nicked site. Newly generated single strands (nicked or nascent strand) are inserted through a groove in Mut L structure. Loop excludes the insertion of double stranded DNA. As DNA is unwound and pushed through Mut L inside the groove, the loop decreases in size. This continues until Uvr D reaches the mismatch. DNA helicase II (Uvr D) and appropriate exonuclease Exo I, Exo VII or Rec J excise error containing DNA beginning at the nick and continuing past the mismatch. DNA polymerase III fills in the gap and DNA ligase seals the remaining nick. The decreased size of loop creates torsional strain that forces the protein-DNA complex to dissociate18.

Mismatch recognition in eukaryotic DNA is done by two heterodimeric complexes of Mut S related proteins-MSH2/GTBP (Mut Sα) and MSH2/MSH3 (Mut Sβ)2,19. Mut Sα binds to both base-base heterologies and small ID heterologies, whereas Mut Sβ plays major role in repair of larger ID mismatches19,20. Presence of ATP markedly decreases the affinity of Mut Sα for an oligonucleotide heteroduplex, an effect also observed with bacterial Mut S. Hydrolysis of ATP facilitates protein-protein interactions and/or sliding along DNA21. Recognition of mismatch is followed by binding of Mut Sα (MSH2/MSH3) with Mut L related proteins [MLH1/hPMS2 (PMS1 in yeast)] and convert into high molecular weight structure. It also increases the efficiency of Mut S proteins to recognize the mismatch. MLH1/MLH3 (PMS2 in humans) also forms the complex with Mut Sβ and help in repair of insertion-deletion mismatches22.

Proliferating cell nuclear antigen (PCNA) has been identified as DNA polymerase processivity factor and has a role in repair at or prior to excision step as a strand discriminating factor12. Bi-directional threading of DNA through hMut S heterodimer is continued till the arrest of forward movement of replication fork that occurs via an interaction with PCNA in the polymerase complex. Recently another eukaryotic mismatch repair endonuclease has been identified whose amino terminal domain is involved in binding to fully methylated DNA and carboxyl terminal region is involved in catalysis and complex formation with Mut L homologue MLH1. This protein is named as MED1 (methyl-CpG binding endonuclease 1). It has got single or double stranded endonuclease activity. Its role in strand discrimination is yet to be identified. Since it shows homology to bacterial damage specific glycosylases/lyases, it is involved in base excision repair.
The properties of MED1 felicitate it to be a functional homologue of Mut H

DNA exonuclease and helicase unwind, nick and degrade the error-containing strand. After dissociation of mismatch repair complex, PCNA, which still is bound at the end of error containing primer strand, recruits replication complex and thus reinitiation starts with the help of DNA polymerase (Fig. 1).

Mismatch repair deficiency, microsatellite instability and tumor development

Defective mismatch repair system, a hallmark of genetic instability can result in unequal crossing over and replication slippage. But now it has been found that mutation in major recombinant genes has no effect on stability. Thus slippage errors, which occur during DNA replication play a major role in causing instability as a result of mismatch repair deficiency. Slippage is more prone in a region of repetitive sequence where DNA polymerase is less able to process through the regions. These repetitive sequences also known as microsatellite sequences are ubiquitously present throughout the human genome. These are short stretch of repetitive DNA e.g. mono- (A)n, di- (CA)n, tri- or tetra- (GATA)n nucleotide repeats where n is usually 5 to 300. In normal mismatch repair proficient cells the number of repeats within a given sequence is stably maintained with accuracy. Microsatellite sequences are generally polymorphic and therefore, used for studies on losses of heterozygosity and linkage analysis. These sequences can also be found in the protein encoding regions of many eukaryotic genes. However, MMR deficient cells show variation at the length of microsatellites which is as a result of slippage error during DNA replication and it can be either insertion/expansion or deletion/contraction of base sequences. This process is termed as microsatellite instability (MSI) also abbreviated as MI or MIN. This phenomenon was originally described in somatic tumor cells compared with the normal tissues and since instability occurs during replication, therefore, it is also known as replication error positive (RER +). Thus, loss of DNA mismatch repair accelerates the process of mutagenesis and result in mutator/microsatellite phenotype. Inactivation of MMR genes also causes instability in coding regions of

Fig. 1—Mechanism for mammalian DNA mismatch repair: (a) Mismatch in daughter strand is introduced as a result of slippage error; (b) Nascent strand containing mismatch is recognized by hMut Sɛ followed by hydrolysis of ATP that facilitates protein-protein interactions or sliding along DNA; (c) Mut Sɛ binds with Mut Lɛ; Bi-directional threading of DNA through hMut Sɛ heterodimer is stopped till it interacts with PCNA as a strand discriminating factor; Pol-ɛ is released; (d) Strand containing mismatch is excised by exonucleases; (e) Resynthesis of DNA daughter strand begins.
transforming growth factor type II receptor (TGFβRII), insulin like growth factor type II receptor (IGFIIIR), hMSH6, hMSH3, BAX genes and contribute to the process of tumourogenesis. It also results in the accumulation of secondary mutations in proto-oncogenes and tumor suppressor genes, thus MMR deficiency leads to cancer development.

Microsatellite instability was first described in colorectal tumours of hereditary non-polypsis colorectal cancer (HNPCC) and later in sporadic colorectal cancers. HNPCC patients develop early onset (40-50 years) synchronous and metachronous tumours and have increased risk for different epithelial tumours including endometrium, stomach, urinary tract, pancreas, and small bowel and ovary tumours. Majority of carcinomas including germline as well as sporadic colorectal cancer display MSI that is a result of defective mismatch repair system. Majority of mutations in HNPCC kinds are found along the entire coding sequence of hMSH2 and hMLH1. In hMSH2, mutations cause premature termination of translational product that is nonsense and frameshift, whereas mutations in hMLH1 are mainly missense and account for 60% of HNPCC patients. Epigenetic changes like hypermethylatation of promoter regions of MMR genes particularly hMLH1 has been found to be associated with the lack of protein expression. This has been observed in sporadic carcinomas like endometrial and gastric tumours. Mutations in hPMS1, hPMS2 and hMSH6 are rare. Chain terminating mutation in hPMS2 brings instability when it is over expressed in MMR proficient cells and thus exhibits dominant negative effect. Moreover, loss of hPMS2 function is compensated by hPMS1. Mutation in amino terminal and carboxy terminal regions of MED1 show dominant negative effects and account for HNPCC and other human carcinomas.

Besides HNPCC, MSI is also observed in extracolonic sporadic tumours belonging to HNPCC spectrum that include endometrial, ovarian, pancreatic, gastric, keratoacanthoma, prostate, breast and other carcinomas. These tumours have been found to harbor multiple genetic alterations at microsatellite sequences and lack MMR gene expression. These tumours exhibit RER+ phenotype to varying degree as 67% in pancreatic tumours to 3% in liver, bladder, brain, ovarian and other gliomas. Different studies investigate loss of heterozygosity and microsatellite alterations as a main genetic defect in urologic malignancies. In other study on squamous cell carcinoma of the head and neck show 45% instability at 2 or more loci and 15% at 40% of the markers studied where as no instability at mononucleotide repeat sequences. Microsatellite alterations in oral squamous cell carcinoma like multiple genetic alterations at 9p21-23 may be indicative of involvement of p16 (CDKN2) tumour suppressor gene at 9p21 in chewing tobacco induced oral cancers.

Testing MSI is one of the non-invasive methods to know about defective MMR system. It may help in the prognosis and tell about the risk of metastasis. Panel of microsatellite markers has been selected which includes mononucleotide repeat motifs-BAT26, BAT40, BAT25, 50C10, 52H10 and DI; dinucleotide CA repeat loci-APC, Mfd15, Mfd26, Mfd28, D3S1283, D9S171, D10S197, D11S1318, D18S58, D18S69 and TP53PCR; trinucleotide repeat loci-HPR1, HPR1II, MYCLI, RB, REN and pentanucleotide repeat loci-FMR2 and TP53 alu. Recently attention has been centered for MSI testing on mononucleotide repeat motifs as a better prognostic marker. Many studies from different populations have shown the quasimonomorphic nature of BAT26 and BAT40 mononucleotide markers. These are very sensitive and highly specific indicators of generalized instability as evident in human cancers. However, in Indian and some other population it has been shown that these loci are polymorphic and therefore, tumoral MSI should be done using corresponding normal DNAs.

**Diagnostic and therapeutic applications**

Detection of MSI has already been recommended in tumor grading of HNPCC. The expression of hMSH2 has been found to be enhanced in low-grade tumors and persists during invasive period. There is one report on the expression of hMSH2 in primary or recurrent tumors in urothelial malignancies. Using urothelial cells in urine the RNA transcript of hMSH2 could be detected in 80% of the patients. Interestingly many were initially negative by cystoscopy and cytology but developed recurrence within surveillance period. MSI has been suggested as a non-invasive tool for early detection and recurrence of bladder carcinoma. MMR deficient or proficient cells respond differentially to many chemotherapeutic agents due to which it has a therapeutic application. MMR deficient cells are more resistant to cisplatinum, busulfan, 5-fluorouracil, mephalan and more sensitive to gamma radiation and may be helpful in predicting tumour response to clinical therapy.
Conclusion

Spontaneous error in DNA replication has been suggested to play a significant role in neoplastic transformation seen in cancer cells. Mutations leading to inactivation of any one of the MMR genes make the genome unstable. Role of mismatch repair and microsatellite instability has been very well established in hereditary non polyposis colorectal cancer but its definite association with many types of sporadic cancers is still debatable. Detection of expression of mismatch repair genes and MSI are valuable tools in developing non invasive markers for cancer.

Acknowledgement

The authors are thankful to CSIR, New Delhi for providing fellowship to Minal Vaish.

References

32. Simpkins S B, Becker T, Swisher E M, Mutch D J, Kovatch A J, Palazzo J P, Fishel R & Goodfellow P J. MLH1 promoter methylation and gene silencing is the primary cause...


