Novel splicing associations of hereditary colon cancer related DNA mismatch repair gene mutations

E Renkonen, H Lohi, H J Järvinen, J-P Mecklin, P Peltomäki

MATERIALS AND METHODS

Patient specimens

Families F73, F95, and F65 meeting the Amsterdam I criteria for hereditary non-polyposis colorectal cancer were screened for predisposing mutations. F73 was part of a population based series, whose clinical and molecular characteristics not related to splicing have been addressed previously. Two HNPCC families with known predisposing mutations in MLH1 (F13 with g→a at nucleotide 454–1 at splice acceptor site of exon 6, and F83 with R659X in exon 17), another two with mutations in MSH2 (F38 with del CA at nucleotide 1550 in exon 10, F25 with ins GT at nucleotide 1860 in exon 12) as well as two healthy control individuals were included in the analysis of splice changes. This study was based on RNA and DNA extracted from Epstein-Barr virus transformed lymphoblasts, supplemented with RNA from uncultured lymphocytes as necessary (the general splice patterns of lymphoblasts and lymphocytes have been shown to be comparable). All human specimens were obtained after informed consent according to the guidelines of the institutional review boards.

RT-PCR and cloning of cDNA

RNA was reverse transcribed with oligo-dT and SuperScript II according to the manufacturer’s instructions (SuperScript First-Strand Synthesis System for RT-PCR, Invitrogen). Full length cDNA of MLH1 (2.3 kb) and MSH2 (2.8 kb) was amplified with primers RT-Lf1 and RT-Lr1 (for MLH1) and RT-Sf1 and RT-Sr1 (for MSH2) (table 1), using Expand Long Template PCR System (Roche). The products were used for direct sequencing to detect mutations as well as for cloning into pCR II-TOPO vector taking advantage of TOPO TA Cloning system (Invitrogen). The resulting plasmids were deposited in the ICG-HNPCC mutation database, mostly affecting MLH1 (≈50%), MSH2 (≈40%), and MSH6 (≈10%) (www.nfdht.nl). Over half of all HNPCC-linked MMR gene mutations consist of nonsense or frameshift changes that result in premature termination codons. Such transcripts are subject to nonsense mediated mRNA decay, a surveillance mechanism whose purpose is to protect the organism against dominant negative or gain of function effects of truncated proteins. Furthermore, some nonsense as well as missense and even silent changes can alter pre-mRNA splicing by introducing or disrupting exonic splicing enhancer or exonic splicing silencer sequences. By doing so, the mutations may promote the skipping or inclusion of exons in which they are located. Evidence in support of this mechanism exists at least for certain nonsense and missense mutations in MLH1. Most mutations that affect splicing consist of single nucleotide substitutions in the classical splice site at intron/exon junctions; such mutations are particularly common in MLH1, constituting one third of all germline mutations in this gene.

In this work, a strategy based on RNA was chosen for mutation screening in families with hereditary non-polyposis colorectal cancer, since a previous study on the same population indicated that 75% of MLH1 and MSH2 mutations were detectable as aberrant sized transcripts. This report focuses on novel splicing associations of the mutations discovered.

Key points

- As part of an effort to identify predisposing mutations in families with hereditary non-polyposis colorectal cancer, full length lymphoblastoid cDNAs of MLH1 and MSH2 were cloned and the products analysed for mutations and splice changes.
- In one family, a nonsense mutation (R100X in MLH1 exon 3) was discovered through its association with skipping of exons not related to the mutation: deletion of one or more exons occurred in 8/14 (57%) among clones originating from the mutated alleles against 3/33 (9%) among clones originating from healthy alleles of two carriers of this mutation (p = 0.001).
- Another mutation affecting the same exon (del AGAA at nucleotide 210) was also associated with multiple exon skipping in a single transcript, but without an overall increase in the frequency of splice events. Five other mutations (two in MLH1 and three in MSH2) were not associated with aberrant splicing.
- Splice events in the mutant MLH1 transcripts clustered around exons 9, 10, and 14–17, which also represent areas involved in alternative splicing.
- Healthy alleles of mutation carriers did not differ from those of healthy control individuals in regard to splicing.
- Our data suggest that some MLH1 transcripts containing mutations are distinguished by an aberrant splicing phenomenon, whose occurrence may depend on mutation location. This splicing may have a biological role, and it may aid mutation diagnostics by distinguishing mutation containing cDNA clones.

Abbreviations: HNPCC, hereditary non-polyposis colorectal cancer; ICG-HNPCC, International Collaborative Group on HNPCC; MLH1, MutL homologue 1; MMR, DNA mismatch repair; MSH2, MutS homologue 2; MSH6, MutS homologue 6; RT-PCR, reverse transcription PCR.
used to transform XL1-Blue (Stratagene, La Jolla, CA) cells, and recombinant clones were identified based on white colour on Xgal containing plates. Among positive colonies, all available or a representative number (average 25) of discrete colonies were selected for DNA isolation and subsequent analysis.

### Analysis of cDNA clones for splicing

To screen for possible size alterations, DNA isolated from individual clones was amplified in three overlapping fragments each for \( \text{MLH1} \) and \( \text{MSH2} \) (table 1) and run through 1.5–2.5% agarose gels. The identity of skipped or inserted exons was determined by sequencing. The allelic origin of transcripts (fig 2) was established using targeted assays for the mutations themselves or their primary splicing consequences, including size fractionation by agarose gel electrophoresis (for \( \text{D}_3, \text{D}_6, \) and \( \text{D}_{17} \)), TaqI (for R100X and R659X) or ApaLI digestion of PCR products (for the \( \text{MSH2} \) exon 10 mutation), and single strand conformation polymorphism analysis (for Q593X). For \( \text{MLH1} \), allelic origin was independently determined taking advantage of a common polymorphism, I219V.

### Studies of circular RNA

RNA was reverse transcribed with random hexamers, and the cDNA was subjected to one round or nested PCR using primers whose orientation made amplification from linear species impossible. To test for circular products involving \( \text{MLH1} \) exons 9 and 10, the following primers were used: 10R, 5'GTTTTTGGGCAAATAGGCTG-3' and 10F, 5'ACACACCACTTCCTGTACC-3'; for a seminested reaction, 9R, 5'GGCAGGATCACTCTGCCC-3' was combined with 10F. The region around \( \text{MLH1} \) exons 14 and 15 was studied with primers 14R, 5'TGCCAAGGCCCACTGAGGATTC-3' and 14F, 5'CTTGGCACAGCATCAAACC-3'; for a seminested reaction, 14R was combined with 15F, 5'-GAAGAACTGTTCTACGATACTC-3'.

![Figure 1](image-url)  
Examples of splice abnormalities observed in cDNA clones from representatives of HNPCC families F73 (F73/1 and F73/2), F95, F13, and F83, as well as a healthy control. All clones representing HNPCC cases correspond to alleles carrying \( \text{MLH1} \) mutations and are marked with superscript letters a–f in this figure as well as in fig 2. For each clone, \( \text{MLH1} \) cDNA was amplified in three overlapping fragments with primers indicated above each panel (see table 1 for primer sequences) and the products were run through 1.5% (left and right panels) or 2.5% (middle panel) agarose gels.

### Table 1

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Direction</th>
<th>Sequence Location</th>
<th>Fragment length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLH1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT-Lf1*</td>
<td>Forward</td>
<td>5'-CATCTAGACGTTTCCTTGCTGCTTCTC-3' 5' UTR</td>
<td>1093 bp</td>
</tr>
<tr>
<td>MLH1-12R</td>
<td>Reverse</td>
<td>5'-GCAAGTCTCCTAGTCAGCAAAGT-3' Exon 12</td>
<td></td>
</tr>
<tr>
<td>MLH1-BF</td>
<td>Forward</td>
<td>5'-AACGGCAGCAGATGCAGG-3' Exon 5</td>
<td>536 bp</td>
</tr>
<tr>
<td>MLH1-1R</td>
<td>Reverse</td>
<td>5'-TCCTAGACGTTTCCTGCTTCTC-3' Exon 11</td>
<td></td>
</tr>
<tr>
<td>MLH1-CF</td>
<td>Forward</td>
<td>5'-CGGAGGACAGACAGACACATG-3' Exon 11</td>
<td>1446 bp</td>
</tr>
<tr>
<td>RT-Lr1*</td>
<td>Reverse</td>
<td>5'-TAAAGGAAATACACTACAGAAGGCAAGTATAA-3' 3' UTR</td>
<td>1084 bp</td>
</tr>
</tbody>
</table>

*From 13

**Online mutation report**

www.jmedgenet.com on April 20, 2022 by guest. Protected by copyright.
Figure 2. Schematic summary of splice changes detected in MLH1 cDNA clones. The allelic transcripts of mutation carriers (figs 2A and B) are primarily distinguished by the presence (fig 2A) or absence (fig 2B) of the predisposing mutation or its primary splicing consequence and secondarily by the base (A or G) present at the polymorphic site in codon 219 (exon 8). F73 (F73/1 and F73/2) and F13, base A at this codon independently identifies the mutant transcripts (those derived from the healthy alleles of the same individuals carry G), whereas F95 and F83 are uninformative due to constitutional homozygosity (both mutant and healthy transcripts carry A). The codon 219 polymorphism defines the allelic origin of transcripts from two healthy controls (figs 2C and D). All splice changes except one (an insertion in a transcript derived from the healthy allele in F13) consist of deletions of one or more exons. The number of clones with a given splice pattern is indicated to the left of each bar. Gel images of splice forms a–f are shown in fig 1. The effect of splice events on the reading frame is given to the right of each bar (parentheses in fig 2A indicate the primary splicing consequence of the predisposing mutation).
RESULTS
Identification of predisposing mutations by cDNA cloning
We originally set out to identify predisposing mutations in three families with hereditary non-polyposis colorectal cancer (F73, F95, and F65). In the first two families, immunohistochemical data (lack of MLH1 protein in tumour tissue) were available, pinpointing the defective gene. Amplification of the 2.5 kb coding region of MLH1 in a single fragment by RT-PCR revealed no aberrant sized fragments by electrophoresis through 1.5% agarose gels (data not shown). Sequencing of RT-PCR products showed a normal sequence in F73, whereas in F95 a double sequence was present, suggesting a heterozygous change. To separate the cDNA products of mutant and healthy alleles, the cDNAs were cloned, and DNA heterozygous change. To separate the cDNA products of mutant and healthy alleles, the cDNAs were cloned, and DNA revealed no aberrant sized fragments by electrophoresis available, pinpointing the defective gene. Amplification of the 2.5 kb coding region of MLH1 in a single fragment by RT-PCR revealed no aberrant sized fragments by electrophoresis through 1.5% agarose gels (data not shown). Sequencing of RT-PCR products showed a normal sequence in F73, whereas in F95 a double sequence was present, suggesting a heterozygous change. To separate the cDNA products of mutant and healthy alleles, the cDNAs were cloned, and DNA

Table 2  Splicing consequences of HNPCC associated MLH1 and MSH2 mutations

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Family</th>
<th>Primary splicing effect</th>
<th>Frequency of mutant clones</th>
<th>Number of clones studied for splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLH1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C→T at n 298, c 100, exon 3 (R100X)</td>
<td>73</td>
<td>None</td>
<td>5/38 (13%)</td>
<td>47</td>
</tr>
<tr>
<td>Del AGAA at n 210, c 70, exon 3</td>
<td>95</td>
<td>Δ3 (in frame)</td>
<td>18/50 (60%)</td>
<td>30</td>
</tr>
<tr>
<td>C→T at n 972, c 659, exon 17 (R659X)</td>
<td>83</td>
<td>Δ17 (in frame)</td>
<td>11/31 (35%)</td>
<td>32</td>
</tr>
<tr>
<td>g→a at n 454-1, splice acceptor of exon 6</td>
<td>13</td>
<td>Δ6 (out of frame)</td>
<td>3/18 (17%)</td>
<td>22</td>
</tr>
<tr>
<td>Healthy control 1</td>
<td>–</td>
<td>–</td>
<td>1/13 (8%)</td>
<td>11 A*</td>
</tr>
<tr>
<td>Healthy control 2</td>
<td>–</td>
<td>–</td>
<td>0/10 (0%)</td>
<td>3/11 B*</td>
</tr>
<tr>
<td><strong>MSH2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C→T at n 1777, c 593, exon 12 (Q593X)</td>
<td>65</td>
<td>None</td>
<td>1/32 (3%)</td>
<td>32</td>
</tr>
<tr>
<td>—with puromycin</td>
<td></td>
<td></td>
<td>2/17 (12%)</td>
<td>17</td>
</tr>
<tr>
<td>Ins GT at n 1860, c 620, exon 12</td>
<td>25</td>
<td>None</td>
<td>3/40 (8%)</td>
<td>40</td>
</tr>
<tr>
<td>Del CA at n 1550, c 518, exon 10</td>
<td>38</td>
<td>None</td>
<td>1/33 (3%)</td>
<td>35</td>
</tr>
<tr>
<td>Healthy control 1</td>
<td>–</td>
<td>–</td>
<td>24</td>
<td>0/1 (0%)</td>
</tr>
</tbody>
</table>

*Among a representative number of clones (18–38) randomly selected for analysis of the predisposing defect.
†Among all clones known to be derived either from the mutation containing or normal allele. Note that even though the frequency of extra splice events in both types of transcripts is unbiased, relative proportion of mutation containing versus normal transcripts is not, since mutation containing clones, which were normally heavily underrepresented, were in some cases actively selected for analysis of splicing.

A similar cDNA based approach turned out unsuccessful in F65. However, genomic exon specific sequencing revealed a novel nonsense mutation, C→T at codon 593, nucleotide 1777, in exon 12 of MSH2, designated as Q593X. This mutation could not be detected in cDNA (with or without cloning) because of severe instability of the respective mRNA: a retrospective single strand conformation polymorphism analysis of 31 originally isolated cDNA clones showed that only one carried the predisposing mutation (see table 2).

Association of DNA mismatch repair gene mutations with aberrant splicing
Prompted by the unexpected splicing consequences of R100X, we included families F73, F95, and F65, together with four additional families with known predisposing mutations (F13 and F83 with MLH1 as well as F25 and F38 with MSH2 mutations) in a detailed analysis of splice changes (table 2, figs 1 and 2). In MLH1, four basic types of splice events were observed. First, deletion of exon 3 in F95 and exon 6 in F13 were compatible with the disruption of a regular exonic and intronic splice site, respectively, through germline mutations. Second, deletion of exon 17 in F83 was likely to reflect the inactivation of an intra-exonic splicing enhancer as a result of the R659X mutation, even though no candidate exonic splicing enhancer sequence could be identified at this site by a prediction program (http://genes.mit.edu/bugelab/rescue-ese/). Third, skipping events in transcripts corresponding to the healthy alleles (the most common being Δ17 and Δ9–10) were compatible with previous RT-PCR reports of alternative splicing. Finally, R100X in F73 was associated with multiple deletion of exons other than those harbouring the predisposing mutation, which constituted a fourth and novel category. Importantly, samples from two affected members of this family (F73/1 and F73/2) both showed the phenomenon: deletion of one or multiple exons occurred in 8/14 (57%) among clones originating from the mutated alleles against 3/33 (9%) among clones originating from the healthy alleles (p = 0.001). While some splice events in the mutant transcripts involved exons that may also be affected by alternative splicing (Δ9, Δ10, Δ9–10, Δ15–16), others affected exons that were never involved in healthy transcripts (Δ14–15, Δ14–16). Unlike mutation containing transcripts, the frequency of splice events in transcripts originating from healthy alleles of these two patients was not elevated as compared to healthy controls. In the entire series, the proportion of transcripts with splice events was similar for healthy alleles from mutation carriers (9/81, 11%) and healthy control individuals (6/44, 14%) (table 2, fig 2).
Our cloning experiments revealed severe underrepresentation of mutant mRNA as a common feature of all three MSH2 mutations investigated, which limited the number of mutant transcripts available for studies of splicing (table 2). In the case of Q993X, translation inhibitor Puromycin was used (200 µg/ml for 6.5 hours) to suppress nonsense mediated mRNA decay, with only a weak effect. No splice events were observed in the mutant MSH2 transcripts. All MSH2 cDNA clones analysed from the healthy control individual also showed the full length product. While the low number of mutant transcripts did not allow comparison between MLH1 and MSH2 in regard to the occurrence of aberrant splicing, splice events in healthy transcripts were less frequent for MSH2 (1/126, 1%) than MLH1 (12/125, 1%) (p<0.01 by χ² test) (table 2, Puromycin experiment not included).

Role of circular RNA formation in exon skipping
To date, a few human genes have been shown to exhibit multiple exon skipping as a result of circular RNA formation.17–21 In this phenomenon, exons are joined at consensus splice sites, but in an order different from that in genomic DNA. Although the mechanism leading to circular splicing products is unknown, it may be provoked by structural aberrations, such as a deletion mutation in the DMD gene.22 To address RNA circularisation in carriers of R100X, a series of RT-PCR experiments was performed on lymphoblastoid and lymphocyte RNA using primers facing away from one other and selecting MLH1 exons 9–10 and 14–15 for our target regions, because of their frequent involvement in splice events (fig 2). By this design, amplification products would be obtained from circular RNA transcripts only, whereas linear RNA species could not serve as templates for PCR. No products consistent with circular RNA were observed (data not shown), thus failing to support this mechanism as an explanation for multiple exon skipping in our case. Our experiments do not exclude the possibility of circular RNA products not involving these target exons.

DISCUSSION
The phenomenon observed for R100X in MLH1—aberrant mRNA splicing of multiple exons unrelated to the mutation—is exceptional and to our knowledge, has not been reported for the MMR genes before. Based on the present relatively limited set of mutations, the occurrence of this abnormality may depend on mutation site and type, perhaps reflecting altered secondary structure of the mutant pre-mRNA, as proposed for another disorder (maple syrup urine disease) displaying multiple exon skipping for an unknown cause.23 The presence of a premature termination codon in itself was clearly insufficient to explain the splicing phenomenon, since among six nonsense or frameshift mutations tested, only R100X showed this associated aberration. Possible local changes in conformation may be compatible with the observation that the only other mutation associated with the skipping of multiple unrelated exons (Δ9 and Δ15), although without an overall increase of splice events, was the 4 bp deletion mutation with the in-frame deletion of the same exon (MLH1 exon 3) as the primary splicing consequence (fig 2).

In rare instances, exons that are skipped may be present as circular RNA products,17–21 and a connection to alternative splicing has been proposed.18 While we did not obtain convincing evidence of RNA circularisation in our study, the regions involved in multiple exon skipping in mutant transcripts overlapped with those commonly affected by alternative splicing in normal human tissues (fig 2)18 19 20. The biological significance of alternative splice products is unknown. Out of frame products are typically unstable (table 2)19, while in-frame products (such as MLH1 Δ9–10, Δ16 and Δ1720) may be otherwise unable to perform their tasks in MMR. More often than restoring the reading frame, altered splicing may induce premature stop codons.2 A recent survey of mRNAs deposited in various databases revealed that 35% of alternative transcripts contained premature termination codons, raising the possibility that alternative splicing connected with nonsense mediated mRNA decay may be a means of regulating protein expression.22 In view of this, it may be important that in our study, 11/15 (73%) of extra splice events observed in mutant MLH1 transcripts were out of frame, as compared to 6/15 (40%) among those occurring in transcripts derived from healthy alleles, even though the difference did not reach statistical significance.

It is presently unknown if the splicing abnormality we describe has any independent influence on clinical manifestations, such as the severity of hereditary non-polyposis colorectal cancer. F73 with R100X showed a relatively late mean age at onset as compared to F95 with in-frame deletion of the same exon and a “milder” form of multiple exon skipping (53 and 46 years, respectively). This difference could reflect the marked decrease of the mutant gene product in the former case, perhaps enhanced by the splicing abnormality, and the shortage of the mutant product could reduce possibilities for dominant negative inhibition.26

All mutations resulting in premature termination codons were associated with a severe reduction in the proportion of the respective transcripts, whereas the allelic transcripts from healthy control individuals were expressed equally (table 2, fig 2). This needs to be taken into account in cDNA based approaches for mutation screening. From the diagnostic point of view, splicing alterations may provide a valuable shortcut to the detection of mutations whose transcripts are too unstable to be visible by simple RT-PCR, making cDNA cloning necessary. As demonstrated for R100X, more than half of clones containing deletions also harboured the predisposing mutation, whereas the clones containing mutations were strongly underrepresented (13%) among all clones. In such instances, instead of random selection, multiple exon deletions may provide a way to single out the likely mutation containing cDNA clones for sequencing.

ACKNOWLEDGEMENTS
We thank Saima Saarinen for expert technical assistance, Jinmin Xiao for cell cultures, Kirsí Pylvänäinen, Tuula Lehtinen, and Heli Surman-Aho for assistance with sample collection, and Tuomo Kukkola-Renkonen for computer support.

Authors’ affiliations
E Renkonen, H Lohi, P Peltomäki, Department of Medical Genetics, FIN-00014 University of Helsinki, Finland
H J Järvinen, Second Department of Surgery, Helsinki University Central Hospital, FIN-00290 Helsinki, Finland
J-P Mecklin, Department of Surgery, Jyväskylä Central Hospital, FIN-40620 Jyväskylä, Finland
This study was supported by the Sigrid Juselius Foundation, the Academy of Finland, the Finnish Cancer Foundation, the Science Foundation of the University of Helsinki, the Paulo Foundation, and the NIH (grant CA82282).

Conflicts of interest: none declared.

Correspondence to: Dr P Peltomäki, Department of Medical Genetics, Biomedicum Helsinki, PO Box 63 (Haartmoninkatu 8), FIN-00014 University of Helsinki, Finland; Paivi.Peltomaki@Helsinki.Fi

REFERENCES

www.jmedgenet.com
Online mutation report