Frequent genomic disorganisation of MLH1 in hereditary non-polyposis colorectal cancer (HNPCC) screened by RT-PCR on puromycin treated samples

I Sumitsuji, K Sugano, T Matsui, N Fukayama, K Yamaguchi, T Akasu, S Fujita, Y Mariya, R Yokoyama, S Nomura, T Yoshida, T Kodama, M Ogawa

Key points

- Three cases of HNPCC with large genomic disorganisations of the MLH1 gene were studied, including two deletions and one duplication.
- RNA was extracted from blood samples after treatment using a translation inhibitor, puromycin, to suppress nonsense mediated mRNA decay. The entire coding region of the MLH1 gene was amplified by reverse transcriptase polymerase chain reaction (RT-PCR) and amplified DNA fragments were subjected to direct sequencing.
- When exon(s) skipping was detected, regardless of the absence of mutations in the corresponding exons or their exon-intron junctions, long distance PCR (LD-PCR) was carried out to amplify relatively large genomic regions and analysed by direct sequencing.
- In the analysis of 28 HNPCC kindreds, seven pathogenic mutations of MLH1 were detected, of which three cases showed large genomic disorganisation, including two genomic deletions and one partial duplication. One of the two deletions was a homologous Alu mediated recombination missing 1222 bp and another was a non-homologous L1 mediated recombination missing 6998 bp with insertion of a guanine residue at the breakpoint. The duplication was considered to be a result of Alu mediated homologous recombination encompassing 20.2 kb from intron 2 to intron 10.
- This study showed that the RT-PCR approach on puromycin treated samples focusing on detection of exon(s) skipping was shown to be a sensitive method to screen large genomic alterations of the MLH1 gene.

MATERIALS AND METHODS

Patients and eligibility criteria

RT-PCR analysis was performed in patients referred to the genetic counselling clinics at the National Cancer Centre Hospital (Tokyo) or Tochigi Cancer Centre Hospital (Utsunomiya, Tochigi). A total of 28 cases were analysed for germline mutations of MSH2 or MLH1 in this study and we focused on three cases that showed large genomic disorganisations of MLH1. All of these cases fulfilled either the Amsterdam criteria II' or the modified Amsterdam criteria.
one should be diagnosed before the age of 50; (5) familial adenomatous polyposis should be excluded in the CRC case(s) if any. The Amsterdam criteria II became adaptable for HNPCC associated cancers, while the original minimum Amsterdam criteria is indicated only for kindreds with colorectal cancer. In Japan, the proportion of CRC cases fulfilling the minimum Amsterdam criteria has been reported to be as low as 0.2%, mostly because of relatively small numbers of family members.15 In our series, 12 cases fulfilled the Amsterdam criteria II, of which 10 cases fulfilled the minimum Amsterdam criteria. In the modified Amsterdam criteria, it is required that (1) very small families, which cannot be further expanded, can be considered to have HNPCC with only two CRCs in first degree relatives, CRC must involve at least two generations, and one CRC case must be diagnosed at <55 years; (2) in families with two first degree relatives affected by CRC, the presence of a third relative with an unusually early onset neoplasm or endometrial cancer is sufficient. Sixteen cases were eligible for the modified Amsterdam criteria and entered into the study.

Blood samples and RNA based analysis

Samples of 20 ml of heparinised peripheral blood lymphocytes (PBL) were obtained from the HNPCC kindreds, who provided informed consent for the study. RNA extraction with or without puromycin, cDNA synthesis, RT-PCR, and direct sequencing reaction were performed as described previously. Briefly, fresh blood samples were incubated in the presence of puromycin (Sigma Chemical Co, MO) at a concentration of 0.2 mg/ml and in the absence of puromycin for two to six hours at 37°C. Leucocytes were isolated from peripheral blood using Vacutainer CPT tubes (Becton Dickinson, NJ). Total RNA was extracted from the leucocytes using the acid guanidium phenol/chloroform method. Reverse transcription was carried out with 200 units of MMLV reverse transcriptase SuperScript II (Life Technologies Inc, MD). Direct sequencing was performed using a Bigdye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems, Inc, CA) and an ABI 310 Genetic Analyzer (PE Applied Biosystems Inc, CA). Primers used for RT-PCR and direct sequencing are shown in fig 1.

Analysis of genomic DNA

DNA extraction was performed by proteinase K digestion and phenol/chloroform extraction as described previously. In cases showing exon(s) skipping in RT-PCR, germline mutations were searched for in the exon-intron junctions that were skipped in RT-PCR (data not shown). When no germline mutations were detected, long distance PCR (LD-PCR) was performed as described previously.

Table 1  Primer sequences used for LD-PCR and breakpoint specific PCR

<table>
<thead>
<tr>
<th>Case No</th>
<th>Name</th>
<th>Sequence</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long PCR primers A</td>
<td>L4-6F1</td>
<td>5'CCTTGTGAGGTGACAGTGGGTGA-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>L4-6R1</td>
<td>5'TGTCCTGGCAAAAGCGAGGTCTTAA-3'</td>
<td>Antisense</td>
</tr>
<tr>
<td>B</td>
<td>L10-3F1</td>
<td>5'CTCAACCAAGACTCACAAGGAACA-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>L10-3R1</td>
<td>5'TGGCTAAATCCTCAAAGGACTGCA-3'</td>
<td>Antisense</td>
</tr>
<tr>
<td>C</td>
<td>Ex1 1F</td>
<td>5'TCCCACTATCTAAGGTAATTG-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>Ex1 1R</td>
<td>5'GGACCAGTCTGAGTGTAAC-3'</td>
<td>Antisense</td>
</tr>
<tr>
<td>PCR primers spanning the mutation breakpoint A</td>
<td>L1 Int 4-5 F2</td>
<td>5'AAATCCAAATCAAGGACTGGA-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>L1 Int 4-5 R2</td>
<td>5'AAATCCAAATCAACCTG TGCC-3'</td>
<td>Antisense</td>
</tr>
<tr>
<td>B</td>
<td>L1 Int 10-2F4</td>
<td>5'CAGGAGGTGCTGATAGTGCTG-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>L1 Int 10-2R3</td>
<td>5'TGGAGATCCTCTCCAGGTGAAC-3'</td>
<td>Antisense</td>
</tr>
<tr>
<td>C</td>
<td>ML1 11F6</td>
<td>5'CAGGAGGAGAAGGATCGATTA-3'</td>
<td>Sense</td>
</tr>
<tr>
<td></td>
<td>ML1 13 R10</td>
<td>5'GGATGCTCCACCCAGAATCCAGTG-3'</td>
<td>Antisense</td>
</tr>
</tbody>
</table>
Figure 2  Gel electrophoresis and direct sequencing profile of RT-PCR products from RNA extracted from PBLs treated with or without puromycin. NP, not treated with puromycin; P, treated with puromycin; M, molecular weight marker λ/HindIII. (A) The arrow indicates the band of approximately 2.4 kb amplified by RT-PCR. (B) Arrows indicate aberrant products of 3.1 kb and normal transcript (2.4 kb). (C) Electrophoresis of the RT-PCR products was carried out using 6% polyacrylamide gels. SV1, SV2, and SV3, three splicing variants; WT, sequencing profile of the normal control. Dotted lines in the sequencing profile indicate exon-exon junctions.
carried out to amplify genomic regions containing the exons skipped in RT-PCR analysis for each patient, using the Advantage Genomic Polymerase Mix (No 8418-1) (Clontech Laboratories Inc, CA). Primer sequences used in LD-PCR are shown in table 1. PCR was performed according to the manufacturer’s recommendations with a minor modification, that is, the final concentration of each primer was 10 µmol/l and reaction mixtures were overlaid with mineral oil. PCR cycling parameters were 94°C for one minute, followed by 40 cycles of shuttle PCR at 94°C for 30 seconds and 68°C for four minutes with a final

Figure 3  Gel electrophoresis of the DNA fragments amplified by LD-PCR from genomic DNA and schematic representation of the genomic rearrangement. M, molecular weight marker λ/HindIII; Pb, proband; WT, wild type. Open boxes indicate exons. Hatched boxes correspond to the Alu elements located in introns and arrows indicate their orientation. The stippled box indicates the particular L1 element included in the breakpoint of case C. Open arrows indicate primer positions for LD-PCR. (A) DNA fragment of approximately 3.2 kb in size was amplified in the proband, while a product of approximately 4.4 kb was obtained from normal genomic DNA. (B) LD-PCR amplified DNA fragment of approximately 3.5 kb from the proband, while no amplification product was obtained from the normal genomic DNA (WT). The black line indicates the presumed direction of homologous recombination, indicating a partial duplication from introns 2 to 10 (boxed region). (C) DNA fragment of about 13 kb was obtained only from the proband in LD-PCR encompassing exons 11 to 14. Schematic representation of the genomic region from exons 11 to 14. The length of the deleted region was 6998 bp indicated by the open box with a single nucleotide “G” inserted at the fusion point.
extension at 68°C for three minutes. Yields of the PCR products were ascertained by electrophoresis on 0.8% agarose gels stained with ethidium bromide.

Sequencing of LD-PCR products was performed as described previously using appropriate primers for each case. Sequences of the cases examined were compared with the human draft sequence for chromosome 3 containing the \textit{MLH1} genomic region (GenBank, accession number NT_005974). The search for repetitive sequences was performed with RepeatMasker and NCBI BLAST analysis.

![Figure 4](http://www.jmedgenet.com/)

**Figure 4** Sequence alignments around the breakpoints. The vertical bars between the sequences indicate positions at which they are identical. The top line and bottom line indicate the normal sequences of introns located 5′ or 3′ of the breakpoints and the middle line indicates the sequence of the breakpoint. (A) The 5′ end of the \textit{Alu} element is indicated by arrows and the core sequences are underlined. Box, 16 bp homologous sequence in introns 4 and 5 that is the presumed region for homologous recombination. (B) These sequences are parts of \textit{Alu} elements and the core sequences are underlined. Box, 44 bp homologous sequence in introns 10 and 2 that is the presumed region for homologous recombination. C, the sequence 5′ of the breakpoint located in the \textit{L1} element. The arrowhead indicates the deletion breakpoint where insertion of a “G” residue is shown by the underlined bold letter.
RESULTS

RT-PCR/direct sequencing

In RT-PCR/direct sequencing approaches, we found three cases showing abnormal exon(s) skipped of MLH1, without mutations in the exon-intron junctions of the corresponding exons (fig 2A, B, C). In case A, the sizes of the RT-PCR products were similar between samples treated with puromycin (P) and without puromycin (NP) (fig 2A), while direct sequencing showed exclusion of exon 5 that was more prominent in P as compared to NP. In case B, gel electrophoresis of the PCR products showed abnormal bands of about 3.1 kb in P and NP, which is larger than the wild type transcript (2.4 kb), and the intensity of the splicing variant was enhanced in P (fig 2B). Sequencing analysis showed variant mRNA representing exon 3 downstream of exon 10, and this was also confirmed by sequencing from the reverse direction in which exon 3 was followed by exon 10, but no exon skipping was observed between exons 10 and 11 (fig 2B). This aberrant mRNA was more abundant in P as compared to NP. In case C, the profiles of the exon skipping were complex and three abnormally spliced DNA bands, splicing variants (SV) 1, SV2 and SV3, were seen on 6% polyacrylamide gel electrophoresis (fig 2C). These bands were excised from the gel, reamplified, and subjected to sequencing. SV1 was shown to skip from exon 11 to 14, while SV2 and SV3 skipped from exon 9 to 14 and from exon 8 to exon 14, respectively.

Long distance PCR (LD-PCR) from genomic DNA

LD-PCR of genomic DNA was performed to amplify the genomic regions including the skipped exons. In case A, showing exclusion of exon 5 (fig 2A), LD-PCR was performed to span genomic regions from introns 3 to 5. The estimated size of the PCR product amplified from a normal allele was 4.4 kb, while a fragment of 3.2 kb was obtained (fig 3A). Direct sequencing of the DNA fragment showed a 1222 bp deletion of the genomic region missing intron 4 to intron 5 (fig 3A). In case B, direct sequencing of the RT-PCR product showed the appearance of exon 10 downstream of exon 3. From this result, we speculated on the presence of a partial duplication flanking exon 3 to 10 in the region downstream of exon 10 (figs 2B and 3B). To confirm this hypothesis, we performed LD-PCR by designing a forward primer in exon 10 and a reverse primer in exon 3 with which normal genomic DNA could not be amplified. A fragment of 3.5 kb was amplified from genomic DNA, while there was no amplification from a control sample (fig 3B). Case C represented a complex form of exon skipping (fig 3C) and we performed LD-PCR using a combination of primer sets designed to amplify genomic regions flanking exons 9-10, 10-11, 11-12, 12-13, 9-11, 10-12, and 11-13, while no differences were detected in comparison with normal genomic DNA (data not shown). Then, we performed LD-PCR spanning 20.1 kb of the genomic region from exons 11 to 14. This showed a fragment of about 13 kb in case C, while no amplification was obtained in normal control DNA. Further primer walking of this DNA fragment showed a genomic deletion of 6998 bp with an insertion of a guanine residue at the breakpoint (fig 3C).

Sequencing analyses of the breakpoints

In all cases, amplified DNA fragments by LD-PCR were sequenced for analysis of the breakpoints (fig 4A, B, C). In cases showing skipping of exon 5 (figs 2A and 3A), the deletion was 1222 bp of which the 5’ breakpoint was located at 415 bp upstream of exon 5, and the 3’ breakpoint was at 733 bp downstream of exon 5 (fig 4A). The breakpoint was located in the Alu element of which the 5’ end was in intron 4 and the 3’ end was in intron 5, joined by a 16 bp homologous sequence in the 5’ sequence of the Alu element (fig 4A). In case B (figs 2B and 3B), sequencing analysis showed that the 5’ breakpoint was located 552 bp downstream of exon 10 and was fused at a point 3002 bp upstream of exon 3. A breakpoint was also located between the Alu elements in introns 10 and 2 joined by 44 bp homologous sequence of which the upstream 19 bp contained the core sequence (fig 4B). In case C, showing genomic deletion from introns 11 to 13 (figs 2C and 3C), the breakpoint was located 1082 bp 5’ of exon 12 and 2621 bp 3’ of exon 13, which caused a deletion of 6998 bp including exons 12 and 13. The 5’ end of the breakpoint was located in the LINE1 (L1) element in intron 11, with a guanine (G) residue inserted at the fusion point and followed by a non-repetitive sequence in intron 13 (fig 4C).

DISCUSSION

Conventional methods for screening MLH1 gene abnormalities are based on exon by exon amplification of the genomic DNA, and such methods are unsuitable for detecting large genomic disorganisation such as large deletions involving whole exons. In this regard, RNA based analyses appeared advantageous, but such methods have a risk of the appearance of normal splicing variants or overlooking deleterious mutations causing NMD. Some investigators concluded that the use of whole blood RNA for screening germline mutations in HNPCC kindreds would be problematical. In this study, we analysed MLH1 and MSH2 by RT-PCR/direct sequencing from RNA samples extracted from PBLS treated with puromycin. The appearance of a few normal splicing isomers was eliminated by use of sequencing primers designed to anneal the neighbouring exons, thus excluding the influence of normal splicing isomers. These analyses allowed us to find deleterious mutations carrying nonsense codons or frameshift changes that are likely to cause NMD and are barely detectable by ordinary RT-PCR approaches. Splicing variants were usually caused by mutations in the exon-intron boundaries or the appearance of a nonsense codon in the skipped exon; however, responsible mutations were not necessarily detected. In this study, we detected three large structural abnormalities of the MLH1 gene in three unrelated families. Cases A and B fulfilled the Amsterdam criteria II, while case C was considered to be a variant form in that the proband suffered from three multiple primary cancers, two colon and one ovarian cancers, but there were only two CRCs in the first degree relatives and both parents died of stomach cancer. In cases A and C, two types of genomic deletion were identified, deletion of 1222 bp and 6998 bp, respectively. In case A, aberrant transcript was identified by direct sequencing, but not by agarose gel electrophoresis of RT-PCR product, in that molecular weights of normal and aberrant transcripts were similar. In case B, a partial duplication between introns 3 and 10 was detected, the size of which reached approximately 20.2 kb in the genomic region. To our knowledge, this is the first report of genomic duplication in MLH1. In cases A and B, exon skipping was obscured in the sequencing profiles of the samples not treated with puromycin (fig 2A, B). Particularly in case A, exon skipping was barely detectable by agarose gel electrophoresis and direct sequencing from the sample without puromycin treatment, elucidating the usefulness of puromycin treatment in detecting deleterious mutations. In cases A and B, no mutations were detected in the corresponding exons or their splice sites, and we suspected that they were caused by large genomic alterations. LD-PCR and subsequent primer walking were carried out to search for the breakpoints. In case A, the breakpoint was located in the left arm of the last Alu element in intron 4 and that of the first Alu element in intron 3, joined by a 16 bp homologous sequence (figs 3A and 4A). In case B, the breakpoint was located in the right arm of the first Alu element in intron 10 and in intron 2 with 44 bp homologous sequences of which the upstream 16 bp belong of the Alu core sequence (fig 4B). Alu mediated homologous recombinations resulted in one large genomic deletion and one partial duplication in these HNPCC kindreds.
In case C, we characterised the large genomic deletions as big as 6998 bp. The 5′ end of the breakpoint was located in the L1 repetitive element in intron 11 and fused 8633 bp upstream of the gene with the insertion of a guanine residue. There were no significant homologies between corresponding normal 5′ and 3′ regions flanking the breakpoint, so non-homologous recombination was a likely mechanism for this deletion. In case C, three types of alternatively spliced bands were detected on RT-PCR analysis (fig 2C). Two of them (SV1 and SV3) showed enhanced intensities in the puromycin treated sample (P), while the second band (SV2) was strong in the sample not treated with puromycin (NP). The pattern of exon skipping in each transcript was different, that is, exons 12 to 13 were excluded in SV1, 10 to 13 in SV2, and 9 to 13 in SV3, respectively. The lengths of the missing exons were 520 bp in SV1 and 881 bp in SV3, while it was 768 bp in SV2.

The differences in signal intensities of these aberrant transcripts with or without puromycin treatment might be the result of the presence or absence of NMD. Skipping of exons 12 to 13 and 9 to 13 resulted in out of frame mutations where nonsense codons appeared 42 bp downstream of exon 14 in SV1 and 19 bp downstream of exon 8 in SV3, respectively. On the other hand, skipping of exons 10 to 13 resulted in an in frame deletion yielding a product missing 256 amino acids in the protein sequence and escaped NMD. These results indicated the effect of puromycin treatment in suppressing NMD and its usefulness for detecting deleterious mutation.

In considerable numbers of HNPPC families, genetic diagnosis is still difficult and approximately 40-50% of the kindreds fulfilling the Amsterdam criteria showed negative results. Large genomic disorganisation may account for these negative results, in part. Searching the ICG-HNPCC mutation database led to the identification of 448 germline mutations of MLH1 and MSH2, respectively (www.nfdht.nl/database/mdbchoice.htm, last update in August 2002), a total of 448 germline mutations of MSH2 and MLH1 have been reported so far (291 for MLH1 and 157 for MSH2) and genomic deletion was found in only 16 cases (3.5%) (eight cases for MLH1 and eight cases for MSH2, respectively). Recently, large genomic deletions of the MSH2 gene were detected in analysis of the rat heterozygote carrying human haploid chromosome from HNPPC kindreds in whom no relevant mutations were detected by the conventional analyses. There were no reports as to the occurrence of large genomic duplications of MLH1 and this study is the first one showing such genomic alterations in the MLH1 gene. Recently, multiplex PCR approaches were introduced in assessing the gene dosage of each exon in either MSH2 or MLH1. This is a useful approach for screening genomic alterations of these genes; however, they did not provide the information for the resulting transcript. As shown in case C, our analysis elucidated that genomic deletion from introns 11 to 13 resulted in three types of splicing variants (fig 2C) and these changes might have implications for understanding the phenotypic manifestation of the case. In our series, a total of 28 kindreds fulfilled either the Amsterdam criteria II or the modified Amsterdam criteria, and pathogenic mutations of either MLH1 or MSH2 were detected in seven and three cases, respectively. Of these, large genomic disorganisation was found in three cases (10.7%), all of which were found only in MLH1, including two large genomic deletions mediated by Alu or L1 elements and one Alu mediated partial duplication. The frequency of large genomic disorganisation in MLH1 reached 43% (3/7) in this study, suggesting that large genomic disorganisation of MLH1 might be a more frequent genetic event. We also indicated that screening for MLH1 germline mutation by RT-PCR/direct sequencing from puromycin treated samples and subsequent analysis using LD-PCR was an efficient diagnostic method for screening for large genomic disorganisations. This approach may also be applicable for diagnosis of other hereditary disorders showing negative results by conventional DNA based screening methods.

ACKNOWLEDGEMENTS

This work was supported in part by Grants in Aid for Cancer Research and for the Second Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare, Japan, and a Grant in Aid from the Vehicle Racing Commemorative Foundation.

Authors' affiliations

K Sumitsuji, K Sugano, T Matsui, N Fukayama, Oncogene Research Unit/ Cancer Prevention Unit, Tochigi Cancer Centre Research Institute, Tochigi, Japan

T Akasu, S Fujita, Y Moriya, Department of Surgery, National Cancer Centre Hospital, Tokyo, Japan

R Yokamoto, Department of Orthopaedics, National Cancer Centre Hospital, Tokyo, Japan

K Sugano, T Yoshida, T Kodama, Genetic Counselling Clinic, National Cancer Centre Hospital, Tokyo, Japan

I Sumitsuji, M Ogawa, Second Department of Surgery, Kumamoto University Medical School, Kumamoto, Japan

K Yamaguchi, Divisions of Clinical Laboratory and Medical Oncology, Saitama Cancer Centre Hospital, Saitama, Japan

S Nomura, Medical Technology Centre, BML Inc, Tokyo, Japan

Correspondence to: Dr K Sugano, Oncogene Research Unit/Cancer Prevention Unit, Tochigi Cancer Centre Research Institute, 4-9-13 Yohnan, Utsunomiya, Tochigi 320-0834, Japan; ksugano@tcp.pref.tochigi.jp

REFERENCES


www.jmedgenet.com
Frequent genomic disorganisation of MLH1 in hereditary non-polyposis colorectal cancer (HNPCC) screened by RT-PCR on puromycin treated samples

I Sumitsuji, K Sugano, T Matsui, N Fukayama, K Yamaguchi, T Akasu, S Fujita, Y Moriya, R Yokoyama, S Nomura, T Yoshida, T Kodama and M Ogawa

J Med Genet 2003 40: e30
doi: 10.1136/jmg.40.3.e30