Identification of Novel Germline hMLH1 Mutations Including a 22 kb Alu-mediated Deletion in Patients with Familial Colorectal Cancer

Jacques L. Mauillon, Pierre Michel, Jean-Marc Limacher, Jean-Baptiste Latouche, Pierre Dechelotte, Françoise Martin, Cosette Martin, Viviane Moreau, Josette Metayer, Bernard Paillot, and Thierry Frebourg

Laboratoire de Génétique Moléculaire [J. L. M., J.-B. L., F. C., C. M., V. M., T. F.], Centre de Dépistage et de Traitement des Tumeurs Digestives [P. M., B. P.], and Laboratoire d’Anatoine et Cytologie Pathologique [J. M.], Centre Hospitalo-Universitaire de Rouen, 1 rue de Germon, 76021 Rouen Cedex; Groupe de Biochimie et de Physiopathologie Digestive et Nutritionnelle, Faculté de Médecine et de Pharmacie, 76800 Saint Etienne dii Rouvray [P. D.]; Hôpitaux Universitaires, 67091 Strasbourg, [J.-M. L.]; and Institut Fédéral de Recherches Multidisciplinaires sur les Peptides, 76821 Mont St Aignan Cedex [J. L. M., P. D., F. C., C. M., V. M., T. F.], France

ABSTRACT

We analyzed the hMLH1 gene in 17 unrelated families with putative hereditary nonpolyposis colorectal cancer. The complete hMLH1 cDNA was amplified in one step, and after a second amplification, four overlapping segments were directly sequenced. We detected, in five families that did not meet the complete Amsterdam criteria, five alterations, including a double-base change resulting in a missense mutation (Lys-618-Ala), a splicing mutation affecting the intron 4 splice acceptor site, a 2-bp deletion at codon 726, a 7-bp deletion at codon 626, and a deletion of exons 13–16. The latter alteration was shown to result from a 22-kb genomic deletion due to a homologous recombination between Alu repeats located in introns 12 and 16. The detection of five germline hMLH1 mutations in five families that only partially fulfilled the Amsterdam criteria shows that these criteria do not allow the identification of all familial colorectal cancers due to mutations of the mismatch repair genes. The numerous Alu repeats present within the hMLH1 gene and the observation of large genomic deletions suggest that Alu-mediated deletions might frequently be involved in hMLH1 inactivation, and (b) reverse transcription-PCR analysis, which allows the amplification of the entire coding region of the hMLH1 gene in one step, might be the most appropriate method for the detection of hMLH1 alterations.

INTRODUCTION

HNPCC, or Lynch Syndrome, is considered the most common form of inherited colorectal cancer (1), and the recent cloning of the HNPCC genes represents, therefore, important progress in the detection of patients with high genetic risk for colorectal cancer. The genes involved in HNPCC are the human homologues of the bacterial MutS and MutL mismatch repair genes (for a review, see Refs. 2 and 3). hMSH2, located on chromosome 2p21, is homologous to MutS (4, 5), and hMLH1 (6, 7), hPMS1, and hPMS2 (8), located on chromosomes 3p21, 2q31-33, and 7p22, respectively, are homologous to MutL. Linkage analysis and mutation reports have suggested that most of the HNPCC cases result from germline mutations of the hMLH1 and hMSH2 genes (9–21).

Detection of a germline mismatch repair gene mutation in a patient with colorectal cancer confirms on a molecular basis the diagnosis of HNPCC, which has important clinical implications for the patient and the patient’s relatives. HNPCC represents a genetic predisposition not only for colorectal cancer but also for a wide spectrum of other neoplasias, such as adenocarcinomas of the endometrium, stomach, ovary, small bowel, and hepatobiliary tract and transitional cell carcinomas of the urinary tract, and numerous patients will develop multiple primary tumors (1, 22). In hMLH1 or hMSH2 mutation carriers, the most important genetic risk is the development of colorectal and endometrial cancers, and the lifetime risks for these cancers have been estimated at approximately 90 and at least 40%, respectively (23). Therefore in carriers, as recommended by the International Collaborative Group on HNPPC, colonoscopy must begin at the age of 25 years and be repeated at least every two years. In women, this surveillance must be completed by annual uterine transvaginal sonography beginning at age 35.

Mutation screening of mismatch repair genes in patients with HNPPC remains difficult for the following reasons: (a) the involvement of four distinct genes; (b) the large distribution of alterations along the coding regions; and (c) the great heterogeneity of the alterations, which include nonsense mutations and small deletions causing premature translation-termination, missense, and splicing mutations. Recently, a 3.5-kb deletion of the hMLH1 gene has been identified in HNPPC families (14). We report here the detection in 17 families of five novel hMLH1 germline mutations, including a 22-kb genomic deletion.

MATERIALS AND METHODS

Patients. This study involved 17 families (Table 1) that fit the following criteria: (a) the family included a patient who developed a colorectal cancer before age 50; (b) the patient had at least one first-degree relative with a colorectal cancer or a tumor belonging to the HNPPC spectrum (22); (c) at least two successive generations were affected; and (d) the family did not exhibit familial adenomatous polyposis. All cancers were histologically verified. Blood samples were collected after informed consent was obtained.

DNA Extraction and RT-PCR. Peripheral blood lymphocytes were purified from 10 ml of blood and placed directly in 400 µl of RNA lysis buffer (Pharmacia Biotech, Uppsala, Sweden) and stored at —20°C. mRNA was extracted, and random hexamer-primed cDNAs were synthesized as described previously (24). The complete open reading frame of hMLH1 was then amplified from 2 µl of the cDNA reaction using MAFS as sense primer and MBRS as antisense primer (Table 2). PCR was performed in a final volume of 20 µl containing 0.5 µM primer and 2.5 units of Pfu DNA polymerase from MBRS as antisense primer (Table 2). PCR products were then cloned into the pUC18 plasmid (Pharmacia Biotech, Uppsala, Sweden) and stored at —20°C. RNA Extraction and RT-PCR. Peripheral blood lymphocytes were purified from 10 ml of blood and placed directly in 400 µl of RNA lysis buffer (Pharmacia Biotech, Uppsala, Sweden) and stored at —20°C. mRNA was extracted, and random hexamer-primed cDNAs were synthesized as described previously (24). The complete open reading frame of hMLH1 was then amplified from 2 µl of the cDNA reaction using MAFS as sense primer and MBRS as antisense primer (Table 2). PCR was performed in a final volume of 20 µl containing 0.5 µM primer and 2.5 units of Pfu DNA polymerase from Stratagene (La Jolla, CA). The PCR consisted of 10 cycles of 15 s at 94°C, 1 min at 60°C, and 2 min at 68°C, then 25 cycles of 15 s at 94°C, 1 min at 60°C, and 2 min plus 20 s added per cycle at 68°C. PCR was preceded by 3 min at 95°C and followed by 5 min at 72°C. The PCR-amplified hMLH1 cDNAs were submitted to electrophoresis on a 1% agarose gel and visualized by ethidium bromide staining.

Sequencing Analysis of hMLH1 cDNA. One µl of the initial PCR was diluted in 100 µl of water, and 1 µl of the diluted template was submitted to a second-stage PCR using four sets of M13 reverse and M13-21-tailed primers (MA1F, MA1R), (MA2F, MA2R), (MB1F, MB1R), and (MB2F, MB2R); Table 2), which generated overlapping segments of 682, 512, 634, and 686 bp, respectively. PCR products were purified by electrophoresis on low-melt agarose gel and directly sequenced on both strands using either the PRISM AmpliTaq or the PRISM AmpliTaq FS Ready Reaction Dye Primer sequencing kits (Applied Biosystems and Perkin-Elmer/Cetus) and an Applied Biosystems model 373A automated sequencer. Identified mutations were confirmed by a second independent RT-PCR analysis. Molecular alterations were further characterized by RT-PCR or amplification of the genomic DNA using primers containing restriction sites. PCR products were then cloned into the...
Bluescript SK(+) vector (Stratagene), and after bacterial transformation, individual clones were sequenced.

**Sequencing Analysis of hMLH1 Introns 12 and 16.** Introns 12 and 16 of the hMLH1 gene were amplified from normal genomic DNA using the 64SF (5'-CGG GAT CCC GGA GAA GAG ACC TAC TTC C 3') and JM2 primers (5'-CGG AAT TCC GGC ACA TCA GAA TCC TCT CG 3') and the 64516F (5'-CGG CAT CCC GAG GAT CTT GGA GAC GAC 3') and 645R (5'-CGG AAT TCC GAT GAG AGG CCC TCC CTC C 3') primers, respectively. The 64SF and 64516F primers contain a BamHI restriction site, and JM2 and 645R an EcoRI restriction site (underlined). To amplify intron 12, we used TagPlus DNA polymerase from Stratagene and an extension time of 3 min. PCR products were then cloned into the Bluescript SK(+) vector, and individual clones were sequenced.

**RESULTS**

We analyzed the hMLH1 gene in 17 families (Table 1) whose presentation was suggestive of the Lynch syndrome, although only 2 met the complete Amsterdam criteria (25). At the time of the study, these families included 46 affected subjects, 12 of whom had developed multiple primary cancers. For each family, we analyzed by RT-PCR the complete open reading frame of the hMLH1 gene were amplified from normal genomic DNA using the 64SF and 64516F primers (5' COO AAT TCC 0CC ACA TCA OAA TCT TCC CO 3') and the 645R (5' COO AAT TCC OAT OAA OAT AGO CAG TCC CTC C 3') primers (Table 1; Fig. 1).

* Family fulfilling the Amsterdam criteria.
* Adenocarcinoma.

<table>
<thead>
<tr>
<th>Family</th>
<th>No. of patients with cancer</th>
<th>No. of cases of colorectal cancer</th>
<th>Other tumors</th>
<th>Age at onset (ys)</th>
<th>Location of the alteration</th>
<th>Nucleotide change</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>2</td>
<td>2</td>
<td>Pharynx</td>
<td>41-44</td>
<td>726-728°</td>
<td>Deletion of CA</td>
</tr>
<tr>
<td>63</td>
<td>2</td>
<td>3</td>
<td></td>
<td>50-65</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>4</td>
<td>2</td>
<td>Endometrium, brain</td>
<td>24-37</td>
<td>626-628°</td>
<td>Deletion of TCTCTTT</td>
</tr>
<tr>
<td>142°</td>
<td>3</td>
<td>4</td>
<td></td>
<td>35-68</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>143</td>
<td>2</td>
<td>2</td>
<td></td>
<td>43-54</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>2</td>
<td>2</td>
<td></td>
<td>41-51</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>174</td>
<td>4</td>
<td>2</td>
<td></td>
<td>47-72</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>189</td>
<td>3</td>
<td>2</td>
<td>Endometrium</td>
<td>42-66</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>2</td>
<td>2</td>
<td></td>
<td>45-50</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>451</td>
<td>4</td>
<td>2</td>
<td>Breast, lung</td>
<td>41-46</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>531</td>
<td>3</td>
<td>4</td>
<td>Bladder, kidney</td>
<td>33-59</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>645</td>
<td>2</td>
<td>2</td>
<td>Endometrium</td>
<td>44-48</td>
<td>Introns 12-16</td>
<td>22.4-kb deletion</td>
</tr>
<tr>
<td>699</td>
<td>1</td>
<td>1</td>
<td>Ovary, small bowel</td>
<td>35-65</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>756</td>
<td>2</td>
<td>2</td>
<td></td>
<td>24-58</td>
<td>618°</td>
<td>AAG→GCG</td>
</tr>
<tr>
<td>910°</td>
<td>4</td>
<td>3</td>
<td>Endometrium</td>
<td>34-54</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>1042</td>
<td>3</td>
<td>1</td>
<td>Breast, ovary</td>
<td>49-68</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>1044</td>
<td>5</td>
<td>2</td>
<td>Ovary, endometrium</td>
<td>32-66</td>
<td>Intron 4 splice acceptor site 5'-tagAGC-3'→5'-tagAGC-3'</td>
<td></td>
</tr>
</tbody>
</table>

In family 24, as shown in Fig. 2, we identified a 2-bp deletion occurring in a CA repeat in exon 19. This deletion, resulting in a stop codon 15 nucleotides downstream, was confirmed by sequencing analysis of genomic DNA (data not shown). In family 73, a 7-bp deletion, resulting in a stop codon 24 nucleotides downstream, was detected in exon 16 by direct sequencing of the cDNA (Fig. 2) and confirmed by cloning and sequencing of the mutant cDNA (data not shown). The proband of family 756 harbored in exon 16 at codon 618 a double base change (AAG→GCG), and cloning of the cDNAs showed that the two nucleotides substitutions were located on the same allele (Fig. 2), resulting, therefore, in a single missense mutation (Lys→Ala). In family 1044, sequencing of the hMLH1 cDNA revealed a deletion of the first 6 bp (AGCAAG) of exon 5 (data not shown). The observation of the AG dinucleotide at the end of the deleted segment, suggesting a mRNA splicing defect, led us to sequence the intron 4-exon 5 boundary in the affected proband. This analysis showed (Fig. 2) a mutation of the splice acceptor site of intron 4 (ag→gg), which therefore resulted in an activation of the cryptic site located at the beginning of exon 5. In family 645, electrophoresis of the PCR-amplified hMLH1 cDNAs revealed an aberrant pattern (Fig. 3a, Lane 1). Amplification of the partial cDNA between exon 12 and 19 confirmed, in addition to the presence of the normal cDNA, the presence of two aberrant shorter cDNAs (Fig. 3b, Lane 1), which were cloned. Sequencing analysis showed that these abnormal cDNAs correspond to a complete deletion of exons 13–16 and exons 13–17. To characterize the deletion at the genomic level, exons 12–17 were amplified from the genomic DNA of the affected proband. A 1.4-kb PCR fragment was observed in the affected subject (Fig. 3b, Lane 3), whereas the normal size of this hMLH1 region is 23.8 kb (10). The sequence revealed that this abnormal genomic PCR product contained exon 12, the 5' end of intron 12, the 3' end of intron 16, and exon 17. To localize the deletion breakpoint, we completely sequenced introns 12 (2.8 kb) and 16 (0.8 kb) of the hMLH1 gene from normal genomic DNA. Analysis of the sequence, using the BLAST program (National Center for Biotechnology Information), revealed that intron 12 contained five Alu repeats, whereas intron 16, as reported previously (14), contained one Alu repeat (Fig. 4a). Alignment of the normal intronic sequences to the sequences observed in family 645 revealed that in this family, the first 570 bp of intron 12 of hMLH1 were fused to the last 670 bp of intron 16 (Fig. 4b). The
Fig. 1. Partial pedigrees of the families with hMLH1 mutation (families 24, 73, 645, 756, and 1044). Filled symbols, affected subjects; open symbols, asymptomatic individuals; oblique lines, deceased; arrow, proband. For each affected subject, the tumor and age of diagnosis are indicated. +/−, heterozygote hMLH1 mutation. Family 1044, shaded symbol corresponds to a mucinous borderline ovarian tumor.

breakpoint (Fig. 4c), which was located within the first Alu repeat of intron 12 and the Alu repeat of intron 16, involved a 32-bp region homologous between introns 12 and 16 (94% homology). These data showed that the 22.4-kb deletion of the hMLH1 gene observed in family 645 was the result of a homologous Alu-mediated recombination and had induced two aberrant splicings between the donor site of intron 12 and the acceptor sites of introns 16 and 17.

DISCUSSION

We analyzed by RT-PCR and direct sequencing the hMLH1 gene in 17 families with familial colorectal cancer, only two of which completely fulfilled the Amsterdam criteria (24) for HNPCC (at least three relatives with histologically verified colorectal cancers, one of whom is a first-degree relative of the other two; at least two successive generations affected; and at least one of the cases of colorectal cancer diagnosed before age 50). We found a germline hMLH1 alteration in 5 of the 15 families that did not meet the complete criteria at the time of the study (33%), and this result is in agreement with the recent report from Nyström-Lahti et al. (18), who observed a germline hMLH1 alteration in 25% of families with HNPCC. These criteria might take into account the size of the family, the development of extra-colonic cancers such as endometrial cancer, and the development of multiple primary tumors, which is a strong indicator of a genetic predisposition for cancer. The five hMLH1 alterations that we describe have not, to our knowledge, been reported previously. Although the size of the families did not allow us to perform a cosegregation analysis between these specific mutations and cancer, we consider that these alterations are probably pathogenic.

The Lys-618-Ala mutation, which results from two nucleotide substitutions, was not found in 28 unrelated subjects, suggesting that this variation is not a common polymorphism. Furthermore, the three lysines present at codons 616—618, which correspond to three AAG repeats, have previously been shown to represent a hot spot for hMLH1 alteration (11, 13, 16, 19, 21). The four other mutations that we reported are also probably pathogenic, because they potentially result in a truncated hMLH1 protein. Two of these mutations highlight two mechanisms of hMLH1 deletion. The CA deletion, observed in family 24, is predicted to result in a truncated protein lacking only the 31 C-terminal amino acids of hMLH1. This CA deletion occurred in...
a sequence of three CA repeats, and a deletion of two CA repeats occurring at the same location was reported previously by Papadopoulos et al. (7). Our report confirms that this dinucleotide repeat, like other di- or trinucleotide repeats characterized within the hMLH1 open reading frame (7, 13, 16, 19, 21), is a hot spot for deletions. It is remarkable to notice that this type of alteration, which is supposed to result from DNA polymerase slippage, will therefore inactivate the hMLH1 gene and lead to a general instability of repeated sequences. In family 645, we described a 22.4-kb deletion, which is the largest deletion of hMLH1 reported thus far. Recently, a 3.5-kb hMLH1 deletion resulting from a recombination between one of the six Alu repeats located in intron 15 and the Alu repeat present in intron 16 has been detected in Finnish HNPCC families (14). The deletion observed in family 645 led us to identify new Alu repeats within hMLH1 and revealed that introns 12 and 16 share a highly homologous 32-bp sequence. Our report and the report from Nystrom-Lahti et al. (14) show that introns of the hMLH1 gene contain numerous Alu repeats, which suggests that large genomic deletions of the hMLH1 gene might

Fig. 2. Detection of hMLH1 mutations in families 24, 73, 756, and 1044. Families 24 and 73, sequence of the RT-PCR product; family 756, sequence of the cloned mutant cDNA; family 1044, sequence of the intron 4-exon 5 boundary amplified from genomic DNA. For each family, the sequence observed in the affected proband (bottom) and the corresponding control sequence (top) are shown. The sequences correspond to the + strand except for family 24. The nucleotide changes are indicated.

Fig. 3. Detection of a 22.4-kb deletion in family 645. a, RT-PCR analysis of the complete hMLH1 cDNA. hMLH1 cDNAs were amplified with the MAFS and MBRS primers (Table 2) as described in “Materials and Methods,” and 10% of the PCR reaction was loaded on a 1% agarose gel. Marker, molecular weight marker BstEII-digested λ; Lane 1, patient 645; Lane 2, normal subject; Lane 3, negative control. The predicted size of the normal PCR product is 2349 bp. b, amplification of the hMLH1 cDNAs between exon 12 and 19 (Lanes 1 and 2) using primers MBlF and MBRS (Table 2) and of the genomic DNA between exon 12 and 17 (Lanes 3 and 4) using primers 64SF and 645R (Table 2). Lane 1, patient 645; Lane 2, normal control; Lane 3, patient 645; Lane 4, normal control; Marker, molecular weight marker HaellIdigested FX174. The predicted size of the normal cDNA is 1190 bp. In patient 645, two other shorter cDNAs were also observed, and sequence analysis revealed that their sizes were 715 and 610 bp. The amplification performed on genomic DNA from patient 645 showed a 1.4-kb product, whereas no amplification product was detected in the normal control. The fragments sizes (in bp) of the markers are indicated.

5731
be frequently involved in \textit{hMLH1} inactivation. These observations have important practical implications, in term of mutations screening. Screening methods of the mismatch repair genes, which are based on PCR amplification of genomic DNA, require that the genes be amplified exon by exon and might not be able to detect large deletions because only the normal allele will selectively be amplified. Although RT-PCR requires specific conditions of samples collection and storage to prevent mRNA degradation, this approach, as shown by this study, allows us to amplify in one step the entire open reading frame of the \textit{hMLH1} gene. The observation of large deletions within the \textit{hMLH1} coding region (this study and Ref. 14) led us to conclude that RT-PCR analysis might be the most appropriate method for the detection and characterization of germ line \textit{hMLH1} alterations.

\section*{ACKNOWLEDGMENTS}

We thank Richard Igo, who gave us the idea to perform RT-PCR with phosphorothioate primers, Jean-Michel Flamant and Benoit Thirion for excellent technical assistance, and Piivi Peltonäki who kindly provided us with the intron 16 sequence of the \textit{hMLH1} gene.

\section*{REFERENCES}

HMLH1 MUTATIONS IN FAMILIAL COLORECTAL CANCER


Identification of Novel Germline hMLH1 Mutations Including a 22 kb Alu-mediated Deletion in Patients with Familial Colorectal Cancer

Jacques L. Mauillon, Pierre Michel, Jean-Marc Limacher, et al.


Updated version  Access the most recent version of this article at: [http://cancerres.aacrjournals.org/content/56/24/5728](http://cancerres.aacrjournals.org/content/56/24/5728)