hMSH2 Mutations in Hereditary Nonpolyposis Colorectal Cancer Kindreds

Bo Liu, Ramon E. Parsons, Stanley R. Hamilton, Gloria M. Petersen, Henry T. Lynch, Patrice Watson, Sanford Markowitz, James K.V. Willson, Jane Green, Albert de la Chapelle, Kenneth W. Kinzler, and Bert Vogelstein

Departments of Oncology (B.L., R. E. P., S. R. H., K. W. K., B. V.) and Pathology (S. R. H.). The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205; The Johns Hopkins University School of Public Health and Hygiene, Baltimore, Maryland 21205 [G. M. P.]; Department of Preventive Medicine and Public Health, Creighton University School of Medicine, Omaha, Nebraska 68178 [H. T. L., P. W.]; Department of Medicine, Ireland Cancer Center, University Hospitals of Cleveland, Case Western Reserve University, Cleveland, Ohio 44106 [S. M., J. K. V. W.]; Memorial University of Newfoundland, St. John's, Newfoundland A1B3V6, Canada [J. G.]; and Department of Medical Genetics, University of Helsinki, FIN-00014 Helsinki, Finland [A. D. L. C.]

Abstract

It has recently been shown that hereditary nonpolyposis colorectal cancer (HNPCC) is caused by hereditable defects in DNA mismatch repair genes. However, the fraction of HNPCC due to defects in any one repair gene and the nature of these mutations are not known. We analyzed 29 HNPCC kindreds for mutations in the prototype DNA mismatch repair gene hMSH2 by a combination of linkage analysis, polymerase chain reaction-based screening, and sequencing of the coding region. The complete intron/exon structure of the gene was ascertained to facilitate this analysis. The results suggest that at least 40% of classic HNPCC kindreds are associated with germline mutations in hMSH2 and that most of these mutations produce drastic alterations in the predicted protein product.

Introduction

HNPCC1 was one of the first familial cancer syndromes described in the biomedical literature (1) but its molecular pathogenesis has only recently become clear. HNPCC appears to be due to hereditary defects in DNA mismatch repair (2–7). In prokaryotes and Saccharomyces cerevisiae, the products of the DNA mismatch repair genes mutS and mutL participate in the recognition and correction of mismatched base pairs resulting from replication errors (reviewed in Ref. 8). These genes have been highly conserved through evolution; five human genes homologous to those responsible for mismatch repair in unicellular organisms have been discovered (2–7, 9, 10). Two of the human genes (hMSH2 and DUG) are related to mutS, while the other three (hMLH1, hPMS1, hPMS2) are homologous to mutL. Mutations in hMSH2 or hMLH1 have been identified in selected HNPCC kindreds.

Cancers that arise in HNPCC patients are genetically unstable (5, 11). It has been shown that nearly all cancers from HNPCC patients (12–14), as well as 12–18% of sporadic colorectal cancers (15–17) and a variable fraction of other tumor types (18–23), have accumulated multiple mutations at repeated sequences distributed throughout their genome. Cell extracts from the tumors of HNPCC patients lack biochemically definable DNA mismatch repair capacity in vitro, while normal cells from HNPCC patients are mismatch repair proficient (5).

On the basis of the biochemical and genetic analyses, it has been suggested that the wild-type allele of the relevant mismatch repair gene is inactivated during tumor formation, thereby increasing the mutation rate and accelerating tumor progression (4). This scenario is formally equivalent to that proposed by Knudson (24) to explain tumorigenesis associated with inherited mutations of tumor suppressor genes.

Although it is currently believed that all HNPCC patients have hereditary defects in DNA mismatch repair genes, the proportion of kindreds with mutations in any specific gene is not known. Presumably, mutations in any of the DNA mismatch repair genes could lead to HNPCC. Information on the nature and number of these mutations is critical for designing effective strategies to detect the mutations in kindreds and to provide appropriate genetic counseling. In this study, we have evaluated a cohort of classically defined HNPCC kindreds to determine the prevalence of mutations in the prototype human DNA mismatch repair gene hMSH2. Our results indicate that at least 40% of classic HNPCC kindreds are caused by mutations in this single gene.

Materials and Methods

Samples. Lymphocytes were obtained from one or more affected members in each kindred. The member(s) chosen for analysis included at least one who was under 50 years of age when diagnosed with colorectal cancer. Two families were from Finland, one was from New Zealand, and the remainder were from North America. RNA and DNA were purified from the fresh lymphocytes or from Epstein-Barr virus-transformed lymphoblastoid cell lines as described previously (25, 26).

Analysis of cDNA. cDNA was generated using random hexamers and RT as described (27). The PCR was used to amplify the hMSH2 transcript in two overlapping fragments. Fragment A contained codons 1–628 of the mRNA and fragment B contained codons 250–934. PCR was performed using 35 cycles of 95°C (30 s), 58°C (1 min) and 70°C (2 min, 30 s) in the buffer described by Sidransky et al. (28). The primers used for RT-PCR included signals for transcription by T7 polymerase and in vitro translation at their 5’ ends (codons 1–628 (fragment A): 5’-GGATCCATAACAGACTACTATGAGGAGACCAATGGCCTGAGCAGAGG-3’ and 5’-CTTGTCTCAAATATGGTTG-3’; codons 250–934 (fragment B): 5’-GGATCCTATACGACCTCACTATGAGGAGACCAATGGCCTGAGCAGAGG-3’ and 5’-GTTATCAATTTACCTCTTTGAGG-3’). Thus, the RT-PCR products could be transcribed and translated in vitro, as described by Powell et al. (27), to search for the presence of mutations which resulted in an altered size of the encoded polypeptide. Controls for RT-PCR included a cDNA sample processed identically except for the omission of RT.

Intron-Exon Borders. The sequences of most intron-exon junctions were ascertained by direct sequencing of the PI genomic clone M1015 (4) using primers chosen from the cDNA sequence (4). M1015 contained all hMSH2 exons except exon 1. A PI clone containing the first exon was obtained by screening a human PI library (Genome Systems, Inc.) with a PCR product corresponding to nucleotides 12–227 of the cDNA. This clone was then used...
Sequencing of independent PCR products from affected patients or their relatives using the primers described in Table 3. The sequence of the exons and flanking intronic sequences were amplified by PCR to define the donor site at the 5' end of intron 1. In selected patients, the sequence of the relevant region of the cDNA was determined by direct sequencing of the RT-PCR products. (d) Mutations were verified by sequencing of genomic PCR products.

The results of this strategy were as follows: (a) Of the 29 families available, linkage analysis was informative in 10 (30–32). Six of the kindreds were shown to be linked to markers surrounding hMSH2 on chromosome 2p16, while linkage was excluded in four families. In the other 19 kindreds, too few individuals were available to make definitive conclusions. (b) and (c) Mutations of hMSH2 were sought in the 6 kindreds linked to chromosome 2p markers and in the 19 kindreds in which chromosome 2p linkage could not be excluded. Ten kindreds with hMSH2 alterations were identified, including seven which have not been reported previously (Table 2). Of these 10, 9 resulted in truncated proteins which could be detected by the IVSP assay. (exam...
Amelogenin PCR products (B and D) were used as templates for direct sequencing as described in “Materials and Methods.” The dideoxyadenosine triphosphate mixes from exhibited in the cDNA of kindred MF (C) (Lane 2, beginning at arrow, Lanes 1 and 3, 3106 (Lanes 2 and 3; Lane 1, normal) but this change is clearly visible in the genomic codon 458 (arrow) is barely visible in the cDNA from two affected members of family each sequencing reaction were loaded in adjacent lanes to facilitate comparison, as were because some of the transcripts encoding truncated proteins were apparently unstable. The best example of this phenomenon was provided by kindred 3106. The presence of a truncated peptide in the RT-PCR products from affected members of this kindred was clear (Fig. 1, Lane 2). Sequence of the resultant cDNA, however, revealed only a very faint band corresponding to a termination codon at residue 458 (TTA to TGA) (Fig. 2A, Lanes 2 and 3). It is known that mRNA encoding truncated proteins is often unstable (33, 34). (d) To confirm and extend the results of the RT-PCR analysis, appropriate exons from genomic DNA were amplified and the products were sequenced in the 10 cases with abnormal cDNA. To enable this genomic analysis, the intron-exon structure of the hMSH2 gene was determined by sequencing PI clones of normal human DNA. The coding region of hMSH2 was thereby found to be divided into 16 exons, each with canonical splice acceptor and donor sites. The sequences of the intron-exon borders are listed in Table 3.

Of the 10 cases with cDNA changes, genomic mutations could be identified in 8. In kindred 3106, for example, a termination codon at residue 458 was clearly observed (Fig. 2B, Lanes 2 and 3). This contrasted with the low intensity of the signal from the RNA product of the mutant allele, as noted above (Fig. 2A). In families C, RB, and MF, the same mutation within the splice donor site at the 3' end of exon 5 was observed, resulting in the deletion of exon 5 from the transcript (Table 2). There was no obvious geographic or ethnic relationship between these three kindreds but we could not exclude shared ancestry. In kindred JG, a similar splice donor mutation (GTT to ATT) at the 3' end of exon 15 was observed, resulting in the deletion of exon 15 in the transcript. In kindred JV, a new donor site was created within exon 12, resulting in the joining of the middle of exon 12 to the splice acceptor site at the 5' end of exon 13. In 2 of the 10 kindreds with abnormal cDNA sequences (TM and DH), we could not identify the genomic changes responsible for the cDNA abnormalities from sequencing of the exons and intron-exon borders. In both of these cases, the changes in the transcript were deletions. It is likely that both of these kindreds had relatively large intragenic deletions within hMSH2. Kindred TM probably had a deletion which included exon 13 and its surrounding intronic sequences, while kindred DH was likely to have a large deletion encompassing exons 8–15.

Discussion

In the 29 kindreds analyzed, 12 (41%) had evidence for involvement of hMSH2. In 10, sequence analysis at the genomic or cDNA levels revealed an alteration of the hMSH2 gene product. In each of those 10, the predicted protein change was substantial, resulting in a large deletion or truncation of the encoded protein (9 cases) or missense mutation in a highly conserved residue (1 case). In two additional families, statistically significant evidence for linkage to hMSH2 was observed, but mutations in the hMSH2 gene could not be found.

These results have substantial implications for genetic diagnosis of HNPCC, especially in relationship to two questions. What is the fraction of HNPCC caused by mutations in hMSH2 and how can these mutations best be detected? In regard to the first question, our data suggest that a minimum of 41% of classic HNPCC kindreds are associated with hMSH2 mutations. The methods we used to screen for mutations would not detect several kinds of mutations that could inactivate the hMSH2 gene. For example, we would not have detected mutations in the promoter, intronic, or 3'-untranslated regions of the hMSH2 gene that resulted in reduced transcription, nor mutations in the coding region that resulted in very unstable transcripts. This point was emphasized by the fact that in two of the six kindreds tightly linked to the hMSH2 locus, we could not identify a mutation in hMSH2, despite sequencing of the entire coding region. Taking our imperfect methods into account and ignoring the possibility that there is another causative gene located close to hMSH2, our data suggest that approximately 50% of classic HNPCC is caused by mutations of this gene. This estimate is consistent with earlier estimates based solely on linkage analysis (12, 31). However, it is difficult to select HNPCC families in a totally unbiased fashion, so larger studies will be needed to confirm this estimate.

The second question concerns the most efficacious way to test for hMSH2 mutations in probands of kindreds with HNPCC. It is apparent from Table 2 that there are few mutational “hot spots,” although three kindreds shared the same mutation. Additionally, the type of mutation seemed to bear little relationship to disease phenotype. Kindreds 3106 and RB, for example, had skin tumors characteristic of the Muir-Torre form of HNPCC, shown previously to be linked to chromosome 2 and associated with microsatellite instability (35, 36), while other kindreds with similar mutations of hMSH2 had no evidence of such skin lesions. It would also appear that evaluation by genomic PCR analysis will not be efficient, when used alone, for detecting mutations in this gene. In kindreds DH and TM, for example, clear alterations of the transcripts were observed on RT-PCR analysis (Table 2) but no mutations could be observed in the genomic DNA. One possible explanation may be that there is large intragenic deletions which included sequences required for amplification of the mutant alleles. Our current strategy for detecting hMSH2 mutations is therefore based on RT-PCR analysis. We first screen these products using the IVSP assay. Nine of the 10 germline mutations of hMSH2 in HNPCC

---

Table 3  Intron-Exon borders

<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon size (base pairs)</th>
<th>Codon&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sequence&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>&gt;279</td>
<td>1–71</td>
<td>5’-GCCTGAGAACATCTTGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>71–122</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>3</td>
<td>279</td>
<td>123–215</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>4</td>
<td>147</td>
<td>216–264</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>5</td>
<td>150</td>
<td>265–314</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>6</td>
<td>134</td>
<td>315–359</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>7</td>
<td>200</td>
<td>359–426</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>8</td>
<td>110</td>
<td>426–462</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>9</td>
<td>124</td>
<td>463–504</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>10</td>
<td>98</td>
<td>505–554</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>11</td>
<td>246</td>
<td>587–669</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>12</td>
<td>205</td>
<td>669–737</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>13</td>
<td>248</td>
<td>737–820</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>14</td>
<td>176</td>
<td>820–878</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
<tr>
<td>16&lt;sup&gt;d&lt;/sup&gt;</td>
<td>&gt;284</td>
<td>879–934</td>
<td>5’-GTAACGAGACGCAGTGGTGGGCGCCTGACGATTCCTTCACCAAAAGGAGGAGGACAGAGGTCAGCTCAGATG</td>
</tr>
</tbody>
</table>

<sup>a</sup> No sequence variations were detected in the samples analyzed other than the mutations listed in Table 2 and a G1A polymorphism of the second position of codon 322.

<sup>b</sup> Upper case letters correspond to exons, lower case to introns. 'Includes untranslated sequences downstream of the initiating ATG (underlined).

<sup>c</sup> Includes 69 base pairs of untranslated sequences upstream of the initiating ATG.

<sup>d</sup> Includes untranslated sequences downstream of the stop codon (underlined).

kindreds thus far observed (Table 2) can be detected by this assay. If a presumptive mutation resulting in an altered size of the polypeptide produced from the RT products cannot be identified, we then recommend single stranded conformation polymorphism, denaturing gradient gel electrophoresis, or direct sequence analysis of either the RT-PCR products or genomic PCR products to detect missense mutations.

In future experiments, we will examine the same cohort of HNPCC kindreds to determine the incidence of mutations in other mismatch repair genes. This should eventually give us a more complete picture of the mutational spectrum in this disease. HNPCC is a relatively common cause of hereditary cancer in western populations, affecting as many as 1 in 200 individuals. Therefore, the study of mismatch repair gene mutations has significant ramifications for the management of the many families afflicted with this disease.

Acknowledgments

The authors thank S. Booker, A. Krush, and J. Cavaliere for clinical coordination and specimen collection and T. Gwiazda for preparing the manuscript.

References

10. Palombo, F., Hughes, M., Jiricny, J., Truong, O., and Hsuon, J. Mismatch repair and
hMSH2 Mutations in Hereditary Nonpolyposis Colorectal Cancer Kindreds

Bo Liu, Ramon E. Parsons, Stanley R. Hamilton, et al.