Germline Mutations in MLH1, MSH2 and MSH6 in Korean Hereditary Non-Polyposis Colorectal Cancer Families

Young-Kyoung Shin,‡ Seung-Chul Heo,‡ Joo-Ho Shin, Sung-Hye Hong, Ja-Lok Ku, Byong-Chul Yoo, Il-Jin Kim, and Jae-Gahb Park

1Korean Hereditary Tumor Registry, Cancer Research Institute and Cancer Research Center, Seoul National University, Seoul, 110-744, Korea; 2Research Institute and Hospital, National Cancer Center, Goyang, Gyeonggi, 411-764, Korea

‡Y.K. Shin and S.C. Heo equally contributed to this work.

Grant sponsors: Research Grant for National Cancer Center; BK21 Project for Medicine, Dentistry, and Pharmacy.

Communicated by Mark H. Paalman

Hereditary non-polyposis colorectal cancer (HNPCC), the most common hereditary colon cancer syndrome, is a dominant disorder caused by germline defects in mismatch repair (MMR) genes. Identification of MMR gene mutations can have direct clinical implications in counseling and management of HNPCC families. We screened 44 HNPCC and 97 suspected HNPCC Korean families for germline mutations in three MMR genes: MLH1, MSH2 and MSH6. We identified twelve novel mutations: nine in MLH1 (c.632_633insT, c.808_811delACTT, c.845C>G, c.1730+1delG, c.1907T>C, c.1930+1delG, c.1918C>T, c.2104-2A>G and c.2170T>A), two in MSH2 (c.1886A>G, c.1316_1318delCCT) and one in MSH6 (c.3488A>T). In addition, two statically significant cSNPs in MLH1: c.1128T>C (p=0.008 in HNPCC and p=0.037 in early-onset CRC) and c.2168C>A (p<0.001 in HNPCC). Interestingly, the most frequent mutation, c.1757_1758insC in MLH1, was a founder mutation inherited from a common Korean ancestor. © 2004 Wiley-Liss, Inc.

KEY WORDS: colorectal cancer; HNPCC; MLH1; MSH2; MSH6; Korean

INTRODUCTION

Hereditary non-polyposis colorectal cancer (HNPCC; MIM# 114500) accounts for ~5% of all colorectal cancer (CRC) cases, making it the most common hereditary colon cancer syndrome [Marra et al., 1995]. HNPCC is characterized by early onset CRC associated with other cancers, including endometrial, small bowel, uterine and...
bladder cancers [Lynch et al., 1999]. HNPCC patients have a tendency to form multiple CRCs, which develop at an average age of 45 years [Lynch et al., 1999].

HNPCC is a dominant disorder associated with germline mutations in mismatch repair (MMR) genes. To date, more than 300 different MMR gene mutations have been identified in approximately 500 HNPCC families [Peltomaki et al., 2001]. HNPCC predisposition has been associated with mutations in MLH1 (MIM# 120436), MSH2 (MIM# 120435), MSH6 (MIM# 600678), PMS2 (MIM# 600259) and PMS1 (MIM# 600258) [Leach et al., 1993; Nicolaides et al., 1994; Papadopoulos et al., 1994], and MLH3 (MIM# 604395) [Wu et al., 2001]. Mutations in MLH1 and MSH2 are thought to occur in about 61% and 36% of HNPCC cases, respectively [International Collaborative Group on HNPCC database, http://www.nfdht.nl].

Predictive genetic testing for germline mutations in MMR genes such as MLH1 or MSH2 can allow identification of HNPCC families, facilitating early treatment and decreasing mortality. Therefore, screening for MMR gene mutations is of direct clinical importance for counseling and management of HNPCC families.

Here, we performed mutational screening of MLH1, MSH2 and MSH6 in 141 Korean families (44 HNPCC and 97 suspected HNPCC). We identified twelve novel germline mutations, a founder mutation and two SNPs significantly associated with HNPCC or early onset colorectal cancer.

**MATERIALS AND METHODS**

**HNPCC families**

To date, 164 HNPCC families (53 HNPCC and 111 suspected HNPCC) have been listed in the Korean Hereditary Tumor Registry. All registered HNPCC families were classified into two subgroups, HNPCC and suspected HNPCC, using the Amsterdam Criteria II and the revised criteria of suspected HNPCC [Park et al., 1999; Park et al., 2002]. The revised criteria for suspected HNPCC are similar to the old suspected criteria I with a few modifications, i.e. the category of ‘vertical transmission of colorectal cancer or at least two siblings affected with colorectal cancer in a family’ was changed to ‘at least two HNPCC associated cancers in first-degree relatives’ [Park et al., 2002]. Among the registered families, 23 (9 HNPCC and 14 suspected HNPCC) had been previously identified as having germline mutations in MMR genes [Han et al., 1995; Han et al., 1996; Yuan et al., 1998; Shin et al., 1999; Shin et al., 2002; http://www.nfdht.nl]; these families were excluded from the present study. We then assessed mutations in three MMR genes, MLH1, MSH2 and MSH6, in the remaining 141 families (44 HNPCC and 97 suspected HNPCC), including newly registered HNPCC families and pedigrees in which mutations had not been previously identified.

For SNP analysis, we also enrolled 42 cases of early-onset CRC (developed before the age of 40 without any family history of disease) collected from Seoul National University Hospital, Korea. Normal population samples were taken from 184 healthy Korean individuals as controls.

**Isolation of genomic DNA**

Peripheral blood lymphocytes were isolated using Ficoll-Paque according to the manufacturer’s instructions (Amersham Biosciences, Uppsala, Sweden). Total genomic DNA was extracted using the TRI reagent following the manufacturer’s instructions (Molecular Research Center, Cincinnati, OH, USA).

**Screening of germline mutations in the MLH1, MSH2 and MSH6 genes**

To identify germline mutations in MLH1, MSH2 and MSH6, we screened all gene coding regions, exon/intron boundaries and core promoters by polymerase chain reaction–single strand conformation polymorphism (PCR-SSCP) analysis. The nucleotide numbers in the mutational notations refer to the wild type cDNA sequence (GenBank NM_000249.2) with the A of the ATG translation initiation start site designated nucleotide +1. The PCR primers, methods and conditions for MLH1, MSH2 and MSH6 gene screening were as previously described [Han et al., 1995; Shin et al., 1999; Shin et al., 2002]. PCR products showing abnormal bands in the PCR-SSCP analysis were bi-directionally sequenced using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and subsequently cloned into pCR2.1 using a TA Cloning Kit (Invitrogen Inc., San Diego, CA, USA) for separate analysis of each allele, and to guard against sequencing errors.
Haplotype analysis of the novel c.1757_1758insC mutation in the MLH1 gene

For haplotyping of individuals with the c.1757_1758insC in the MLH1 gene, 13 highly polymorphic microsatellite markers from chromosome 3p were analyzed (D3S1283, D3S1266, D3S1609, D3S1561, D3S3623, D3S3718, D3S3512, D3S1277, D3S1612, D3S1611, D3S1298, D3S1260, D3S3564) [Nystrom-Lahti et al., 1994; Dip et al., 1996; Jager et al., 1997; Collins et al., 1998]. Genotyping was performed using marker sequences from the Genome Database (www.gdb.org). All forward primers were tagged with M13(-21) [Schuelke et al., 2000]. PCR was performed in the presence of fluorescent dye-labeled M13(-21) primers, and amplified products were analyzed on an ABI 3100 sequencer using the GeneScan Software 3.7 and GenoTyper 2.0 (Applied Biosystems, Foster City, CA, USA).

Statistical analysis

Comparisons of the allele frequencies between controls and cases were performed using the $\chi^2$ test or Fisher's exact test with SPSS software (version 9.0, SPSS Inc, Illinois, USA). Odds ratio (OR) and 95% confidence intervals (CI) were also calculated using this software. A level of $p<0.05$ was considered statistically significant.

RESULTS

Germline mutations in the MLH1, MSH2 and MSH6 genes

Out of 44 HNPCC and 97 suspected HNPCC families, we identified 21 families with germline mutations in the tested MMR genes (Table 1). Of these, twelve families each had one novel germline mutation (bold in Table 1), while nine families carried three previously reported germline mutations. Of the twelve novel germline mutations, nine were found in MLH1, two in MSH2 and one in MSH6. The nine novel MLH1 germline mutations included two frameshift mutations (c.622_623insT and c.808_811delACTT), two splicing defects (c.2104-2A>G and c.1730+1delG) and five missense mutations (c.845C>G, p.Ala282Gly; c.1625A>C, p.Gln542Pro; c.1907T>C, p.Leu636Pro; c.1918C>T, p.Pro640Ser; c.2170T>A, p.Leu724Met). The two novel MSH2 mutations were a c.1886A>G (p.Gln629Arg) missense and a c.1316_1318delCCT in-frame deletion. In MSH6, we found a single novel c.3488A>T (p.Glu1163Val) missense mutation.

Of the nine families with three previously reported germline mutations, one (SNU-H3) had a c.1721T>C (p.Leu574Pro) mutation in exon 15 of MLH1, six (SNU-H10, SNU-H30, SNU-H35, SNU-H1035, SNU-H1070, NCC-H1004) had c.1757_1758insC mutations in exon 16 of MLH1, and two (SNU-H5005, SNU-H1069) had G-to-C transversion mutations at position –225 in MSH2.

The Korean Hereditary Tumor Registry contains 164 HNPCC families (53 HNPCC and 111 suspected HNPCC). Together, our work and previous reports on 9 HNPCC and 14 suspected HNPCC families [Han et al., 1995; Han et al., 1996; Yuan et al., 1998; Shin et al., 1999; Shin et al., 2002; http://www.nfdht.nl] indicate that germline mutations are present in the MLH1, MSH2 or MSH6 genes of 22 of the 53 HNPCC families (41.5%; 50% in MLH1, 4% in MSH2 and 4% in MSH6), and 22 of the 111 suspected HNPCC families (19.8% ; 50% in MLH1, 41% in MSH2, 9% in MSH6).

Association between cSNPs and HNPCC

We identified four coding single nucleotide polymorphisms (cSNPs), three (c.1128T>C, c.1151T>A and c.2168C>A) in MLH1 and one (c.471C>A) in MSH2. Three of these cSNPs are novel polymorphisms, while c.1157T>A had been previously reported [Kobayashi et al., 1996]. We investigated whether these cSNPs were associated with HNPCC and/or early-onset CRC. The c.1128T>C polymorphism was found to occur more frequently in both HNPCC ($p=0.008$, OR=3.3, 95% CI: 1.30 ~ 8.35) and early-onset CRC ($p=0.037$, OR=2.7, 95% CI: 1.03 ~ 7.07) patients as compared to controls; however, there was no statistically significant difference in c.1128T>C frequency between suspected HNPCC patients and controls. The c.2168C>A variant was positively associated with HNPCC ($p<0.001$, OR=11.7, 95% CI: 2.95 ~ 46.2) patients, but not with suspected HNPCC or early-onset CRC patients. The c.471C>A polymorphism in the MSH2 gene showed no significant association with HNPCC, suspected HNPCC or early-onset CRC patients. Families harboring the c.2168C>A cSNP did not show mutations in MLH1, MSH2 or MSH6, whereas 9 out of the 18 families with c.1128T>C (Asp376) also had pathogenic mutations in MLH1, MSH2 or MSH6.
Twelve families in bold each had one novel germline mutation.
c.1757_1758insC - a founder mutation in MLH1

Eleven (35%) of the 31 families with MLH1 mutations harbored a c.1757_1758insC mutation. The genotypes of 52 individuals from these 11 families were analyzed using a set of 13 highly polymorphic microsatellite markers from chromosome 3p. All families carrying this mutation were found to harbor a single allele for the intragenic marker, D3S1611, which was mostly absent in the unaffected chromosome. In addition, these individuals all shared portions of a 10 marker haplotype around MLH1 (Fig. 1). In the disease chromosome, marked linkage disequilibrium was observed for 10 of the 13 tested markers, but not for the telomeric markers (D3S1283 and D3S1266) or the flanking marker (D3S1609). These data indicate that all 11 families with the MLH1 c.1757_1758insC mutation inherited this mutation from a common ancestor.

**Table 2. SNP Genotypes and Allele Frequencies (Cases vs. Controls)**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Position</th>
<th>n</th>
<th>Genotypes</th>
<th>Allele Frequency</th>
<th>p-value (Odds Ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLH1</td>
<td>c.1128T&gt;C (Asp376)</td>
<td></td>
<td>T/T</td>
<td>C/C</td>
<td>T</td>
</tr>
<tr>
<td>Control</td>
<td>184</td>
<td>172</td>
<td>12</td>
<td>0</td>
<td>356 (0.97)</td>
</tr>
<tr>
<td>HNPCC</td>
<td>40</td>
<td>32</td>
<td>8</td>
<td>0</td>
<td>72 (0.90)</td>
</tr>
<tr>
<td>shHNPCC</td>
<td>72</td>
<td>62</td>
<td>10</td>
<td>0</td>
<td>134 (0.93)</td>
</tr>
<tr>
<td>Early-onset</td>
<td>42</td>
<td>35</td>
<td>7</td>
<td>0</td>
<td>77 (0.92)</td>
</tr>
<tr>
<td>c.2168C&gt;A (p.Ala723Asp)</td>
<td></td>
<td></td>
<td>C/C</td>
<td>C/A</td>
<td>A</td>
</tr>
<tr>
<td>Control</td>
<td>184</td>
<td>181</td>
<td>3</td>
<td>0</td>
<td>365 (0.99)</td>
</tr>
<tr>
<td>HNPCC</td>
<td>40</td>
<td>33</td>
<td>7</td>
<td>0</td>
<td>73 (0.91)</td>
</tr>
<tr>
<td>shHNPCC</td>
<td>72</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>144 (1.00)</td>
</tr>
<tr>
<td>Early-onset</td>
<td>42</td>
<td>42</td>
<td>0</td>
<td>0</td>
<td>84 (1.00)</td>
</tr>
</tbody>
</table>

Numbers in the n column indicate the number of individuals. In the allele frequency column, numbers show the number of chromosomes and parentheses show allele frequency. shHNPCC, suspected HNPCC; CRC, colorectal cancer.

*: Where statistical significance may be considered borderline, OR (Odds Ratio) is also shown.

**: Using Fisher’s exact test.

**Figure 1.** Disease haplotype of 11 families with the MLH1 c.1757_1758insC mutation. Markers are listed in linear order at left. The conserved haplotype is shown in bold face.

**DISCUSSION**

Our group has previously identified 23 MMR germline gene mutations in Korean HNPCC families [Han et al., 1995; Han et al., 1996; Yuan et al., 1998; Shin et al., 1999; Shin et al., 2002; http://www.nfdht.nl]; here, we identified 12 novel germline mutations in the MLH1, MSH2 and MSH6 genes. These mutations were present in 22 (41.5%) of 53 HNPCC families contained within the Korean Hereditary Tumor Registry, and 22 (19.8%) of 111 suspected HNPCC families in the Registry.
Of the 12 novel germline mutations identified in the present study, 11 are predicted to cause abnormalities in the MMR gene product. Two of the identified MLH1 mutations, c.632_633insT and c.808_811delACTT, are frameshift mutations predicted to result in truncated protein products. Two other MLH1 mutations, c.1730+1delG and c.2104-2A>G, are predicted to result in splicing defects in exons 15 and 19, respectively. Four missense mutations, p.Gln542Pro (c.1625A>C), p.Leu636Pro (c.1907T>C), p.Pro640Ser (c.1918G>T) and p.Leu724Met (c.2170T>A), are predicted to lead to amino acid polarity changes in the PMS2 binding domain of the MLH1 gene product, suggesting they would affect MLH1 protein function. The p.Gln629Arg (c.1886A>G) mutation is located in the MLH1 functional domain and would therefore be likely to hinder protein interactions. Though the observed p.Glu1163Val (c.3488A>T) mutation in exon 6 of MSH6 is not located in a functional domain, the substitution of a negatively charged amino acid to a non-polar amino acid may cause abnormal MSH6 protein function, as might the observed MSH2 in-frame deletion, c.1316_1318delCCT, which causes the deletion of one amino acid (Leucine).

As shown in Table 1, we identified a total of 31 MLH1 mutations in the screened families. Most of these mutations are predicted to cause truncated MLH1 proteins (54.8%, 17/31). The tested Korean HNPCC families showed a high frequency of the MLH1 c.1757_1758insC mutation, contributing to an MLH1 mutation rate of 90%, which is more than 20 times the MSH2 mutation rate (4.5%) in HNPCC families. The c.1757_1758insC mutation in MLH1 accounted for 35% of MLH1 mutations in HNPCC and suspected HNPCC families; it was present in 11% of HNPCC and 4.5% of suspected HNPCC families. We further tested whether there might be a founder effect at work in Korean HNPCC families.

Haplotype analysis revealed that the MLH1 c.1757_1758insC mutation can be associated with a single ancestral founding mutation in Korean HNPCC families. Interestingly, historical analysis revealed that all 11 families segregating this mutation originated from the southern part of the Korean peninsula (data not shown). The high proportion of c.1757_1758insC mutations in Korean HNPCC families indicates that screening for the MLH1 c.1757_1758insC mutation in this population could be very helpful in clinical counseling.

HNPCC patients are at risk of developing CRC (80%), endometrial cancer (40% for women), gastric cancer (15%) and ovarian cancer (10% for women) [Arnio et al., 1995; Lynch and Smyrk, 1996]. However, our previous data indicated that Korean HNPCC patients had a considerably lower risk (4.7%) of endometrial cancer [Park et al., 1999]. Interestingly, endometrial cancer was present in 4 of 36 women in the 19 families with MLH1 mutations other than c.1757_1758insC, but was not observed in any women (n=19) from the 11 families with the MLH1 c.1757_1758insC mutation. The mean age at final follow-up for the 19 women with the c.1757_1758insC mutation was 49.4 years, whereas the mean age of endometrial cancer onset in Korean women with HNPCC is 42.3 years [Park et al., 1999]. The apparent relevance of the MLH1 c.1757_1758insC mutation to endometrial cancer risk demonstrates the importance of genetic testing for the effective clinical management of Korean women with HNPCC.

In this study, we identified three novel SNPs in MLH1 and MSH2. Of these, the c.1128T>C polymorphism was significantly more frequent in both HNPCC (P=0.008, OR=3.3, 95% confidence intervals: 1.30–8.35) and early-onset CRC patients (P=0.037, OR=2.7, 95% CI: 1.03–7.07), as compared to the controls. The c.2168C>A variant was also significantly associated with HNPCC (P<0.001, OR=11.7, 95% CI: 2.95–46.2), but not with suspected HNPCC or early-onset colorectal cancer. In several studies, germline mutations have been considered the most significant risk factor for the development of HNPCC. However, ~33% of HNPCC cases have been shown not to harbor germline mutations in any tested MMR gene [Peltomaki, 2001]. It is interesting to speculate on what other factors might cause HNPCC susceptibility. As a result of attempting to elucidate these other factors, a number of polymorphisms have been found in genes associated with HNPCC, and it was suggested that the causative defects appear to lie in other as yet unidentified genes [Boland, 2000]. SNPs can be located in both coding and regulatory regions, and can profoundly affect the function of the encoded gene product [Collins et al., 1998]. For example, the c.2168C>A in MLH1, which causes the substitution of a negatively charged amino acid with a non-polar amino acid, may cause abnormal MLH1 protein function.

In summary, a total of twenty-one germline mutations and four SNPs were identified in three MMR genes, MLH1, MSH2 and MSH6, in 141 Korean HNPCC families (44 HNPCC and 97 suspected HNPCC). Of these, twelve mutations (nine MLH1 mutations, two MSH2 mutations and one MSH6 mutation) and three cSNPs in MLH1 were identified as novel. The MLH1 c.1757_1758insC mutation was identified as a founder mutation inherited from a common ancestor. Although further functional studies will be necessary to determine how each of these MMR gene mutations affect the protein functions, our observation that colon cancer segregates with these
mutations in most of the screened families indicates that predictive testing and genetic counseling should be strongly considered.

REFERENCES


