Identification and Characterization of Genomic Rearrangements of MSH2 and MLH1 in Lynch Syndrome (HNPCC) by Novel Techniques

Hidewaki Nakagawa, Heather Hampel, and Albert de la Chapelle*

Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio

*Correspondence to: Albert de la Chapelle, Human Cancer Genetics, 646 Tzagournis Medical Research Facility, 420 W. 12th Avenue, Columbus, OH 43210. E-mail: delachapelle-1@medctr.osu.edu

Grant sponsor: National Institutes of Health; Grant numbers: CA67941, CA16058

Communicated by Daniel Schorderet

It has recently been suggested that large genomic rearrangements account for 10-20% of all MSH2 mutations, and a lower proportion of all MLH1 mutations, among individuals with Lynch syndrome (hereditary non-polyposis colorectal cancer, HNPCC). These rearrangements are notoriously difficult to detect; moreover, for clinical purposes, simple tests must be devised to screen family members at risk. Here we used the multiplex ligation-dependent probe amplification (MLPA) method to screen for MSH2 and MLH1 deletions in 70 patients whose colorectal or endometrial tumors were MSI positive, yet no mutation had been found by genomic exon-by-exon sequencing of MSH2, MLH1, and MSH6. We identified five candidates with four different MSH2 deletions (exons 1-2, exons 1-6, exons 1-7 and exon 8) and one candidate with an MLH1 deletion (exons 3-6). To confirm the screening results and to characterize the breakpoints of these genomic deletions precisely, we used diploid-to-haploid conversion and inverse PCR as well as long-range PCR. In each case, we were able to pinpoint the breakpoint and design a simple diagnostic PCR. The procedures we used appear to be sensitive, specific, and simple enough for clinical use. © 2003 Wiley-Liss, Inc.

KEY WORDS: Lynch syndrome; HNPCC; MSH2; MLH1; genomic deletion; MLPA

INTRODUCTION

Lynch syndrome (LS), also known as hereditary nonpolyposis colorectal cancer (HNPCC, MIM# 114500) accounts for 1-5% of all colorectal cancers, and an unknown proportion of endometrial, gastric, ovarian, and several other cancers (reviewed in Lynch and de la Chapelle, 2003). Germline mutations of several mismatch repair genes cause LS (Peltomäki and Vasen, 1997; Papadopolous and Lindblom, 1997). Until recently, only very few large genomic rearrangements had been described in these genes, e.g. a deletion of exon 16 of MLH1 (MIM# 120436), that is a founder mutation in Finland (Nyström-Lahti et al., 1995). The existence of many deletions and other rearrangements in these genes was long ignored because the most commonly used mutation detection methods (e.g. exon-by-exon sequencing or SSCP) by definition are not able to detect such abnormalities. By Southern hybridization Wijnen et al. first demonstrated that genomic deletions of MSH2 (MIM# 120435) were much more prevalent than previously thought (Wijnen et al., 1998). Subsequently, several PCR-based
semiquantitative DNA dosage measurement techniques have been developed (Armour et al., 2002) and the relatively high prevalence of large genomic rearrangements in MSH2 confirmed (Charbonnier et al., 2002; Gille et al., 2002; Wang et al., 2003). Moreover, large genomic rearrangements of MLH1 are also relatively common (Wang et al., 2002; Nakagawa et al., 2002; Gille et al., 2002).

There is a clear clinical need for simple and reliable means of screening for these rearrangements. Importantly, for each new molecular rearrangement thus detected, it is desirable to devise a simple PCR-based diagnostic method to search for the mutation in at-risk family members of the proband. This usually requires that the breakpoints of the rearrangement are precisely determined. So far this has only rarely been reported (Charbonnier et al., 2002; Gille et al., 2002; Wang et al., 2003).

In this study we screened germline DNA of 70 individuals at risk for LS by multiplex ligation-dependent probe amplification (MLPA) and found 6 presumptive deletion cases. These were then confirmed and characterized by using novel strategies and in each case, a simple PCR reaction was devised as a diagnostic tool.

**MATERIALS AND METHODS**

**Patients**

We selected 57 consenting patients with newly diagnosed MSI-positive colorectal cancer (n=38) or endometrial cancer (n=19) from an on-going cohort study in Columbus, Ohio, and 13 patients from a series of high-risk HNPCC families counseled by the Clinical Cancer Genetics Program at the James Cancer Hospital, the Ohio State University. Most of the 57 cohort study patients did not have clinical or family features highly suggestive of LS. Among all 70 patients only 11 were positive for the Amsterdam 1 or 2 criteria, while an additional 16 were positive for the Bethesda criteria (Lynch and de la Chapelle, 2003). MSI status was evaluated by BAT25, BAT26, D18S69, D2S123, and D5S46. In all cases, deleterious point mutations of MLH, MSH2 and MSH6 had been excluded by exon-by-exon sequencing. All experiments were performed on genomic DNA from blood, lymphoblastoid cell cultures, or mouse-human hybrid cell clones.

Multiplex ligation-dependent probe amplification (MLPA). MLPA as described by Schouten et al. (2002) was used. A MLPA test kit for MSH2 and MLH1 was obtained from MRC-Holland, Amsterdam, The Netherlands (Gille et al., 2002). The probe mixture contained 16 exon probe pairs for MSH2 and 19 exon probe pairs for MLH1, plus seven control-probe pairs. MLPA was performed according to the instructions. Briefly, 100 ng genomic DNA was denatured at 98°C for 5 min and hybridized with probe mix in a tube at 60°C overnight. Ligation was performed using heat-stable Ligase-65 at 55°C 15min. Next PCR was performed by using one FAM-labeled primer and one unlabeled primer for 35 cycles of 95°C 30s, 60°C 30s and 72°C 1min. The FAM-labeled PCR product was analyzed on an ABI 3700 sequencer and by Genotyper software. Deletion was suspected when the peak height was 60% or less of normal controls.

**Conversion to haploidy**

The diploid-to-haploid conversion technique was described previously (Yan et al., 2000, Nakagawa et al., 2002). Haploid converted clones from our patients were created by GMP Genetics, Inc, Boston. In brief, human fresh lymphocytes or immortalized lymphoblastoid cells were electrofused with a specifically designed mouse cell line (E2). Unfused mouse parental cells were negatively selected by HAT (Gibco BRL) and unfused human cells were negatively selected by Geneticin (Gibco BRL).

**Inverse PCR and sequencing**

Inverse PCR is a convenient and rapid method to identify unknown sequences following known sequences. It was performed as previously described (Wagner et al., 2002). Five µg haploid genomic DNA from mouse hybrid cells containing the mutant allele of human chromosome 2 was digested by several appropriate restriction enzymes overnight. After purification, 1 µg digested DNA was self-ligated in 400ul 1x T4 ligase buffer overnight at 14°C. Nested PCR amplified the target region by using exon 3-specific primers (case 2), intron 6-specific primers (cases 3 and 4) and intron 7-specific primers (case 5). The amplified PCR product was TA cloned and sequenced on ABI 3700.

**Deletion-specific PCR spanning the breakpoints**

Deletion-specific PCR was designed by using the Expand Long Template PCR system (Roche) according to the instructions. Primers and the product size of each deletion specific PCR are shown in Table 1.
MSH2 and MLH1 Genomic Rearrangements

Table 1. Characterization of Six Genomic Deletions of MSH2 and MLH1 in Lynch Syndrome

<table>
<thead>
<tr>
<th>No</th>
<th>Gene</th>
<th>Del exons</th>
<th>Deletion designation</th>
<th>Primers of deletion-specific PCR</th>
<th>Product size</th>
<th>Clinical criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MSH2</td>
<td>8</td>
<td>g.26488048_26491055del3008 (NT_022184)</td>
<td>5'-tgagttggaattcagttgcgtcgt-3’ 5'-ttatacctacacacggtgcacac-3’ 5'-gcactcagttgaacaaaaggtggtggcag-3’ 5'-gctgacatatcattgttaccaaagaga-3’</td>
<td>1.3kb</td>
<td>Bethesda</td>
</tr>
<tr>
<td>2</td>
<td>MSH2</td>
<td>1-2</td>
<td>g.26441535_26452813del11279 (NT_022184)</td>
<td>5'-aagcatcacagttactgttg-3’ 5'-gctgaattaggttttggaac-3’ 5'-cacagacctaggatcaagatttatggc-3’ 5'-caccagagaagtacaccttttaaagtt-3’</td>
<td>0.5 kb</td>
<td>Bethesda</td>
</tr>
<tr>
<td>3</td>
<td>MSH2</td>
<td>1-6</td>
<td>g.26445441_26465485del20045 (NT_022184)</td>
<td>5'-aggcctgaagttgattcagatccaa-3’ 5'-aactgagaaactaatgcctgcattgtgta-3’</td>
<td>1.7 kb</td>
<td>Amsterdam2</td>
</tr>
<tr>
<td>4</td>
<td>MSH2</td>
<td>1-6</td>
<td>g.26445441_26465485del20045 (NT_022184)</td>
<td>5'-aggtcaaatcagttaggtggcagatccaa-3’ 5'-aactgagaaactaatgcctgcattgtgta-3’</td>
<td>1.7 kb</td>
<td>Bethesda</td>
</tr>
<tr>
<td>5</td>
<td>MSH2</td>
<td>1-7</td>
<td>g.26399108_26483640del184533 (NT_022184)</td>
<td>5'-aagtgaataatcctcctgtggagttggtg-3’ 5'-aactgagaaactaatgcctgcattgtgta-3’</td>
<td>1.7 kb</td>
<td>Bethesda</td>
</tr>
<tr>
<td>6</td>
<td>MLH1</td>
<td>3-6</td>
<td>g.18552341_18564985del12645 (NT_022184)</td>
<td>5'-aagtgaataatcctcctgtggagttggtg-3’</td>
<td>2.0 kb</td>
<td>Amsterdam1</td>
</tr>
</tbody>
</table>

a) GenBank: AH003235, NT_022184
b) from Wagner et al., 2003.
c) GenBank: AH003234, NT_022517

RESULTS

Among the 70 selected patients with MSI-positive colorectal or endometrial cancer and/or with clinical risk of LS we found an abnormal chromatogram pattern in ten patients by MLPA assay. Six of these patients showed consistently normal control peaks as described by the kit’s manufacturer, and chromatogram patterns that were tentatively interpreted to indicate genomic deletions involving exons of MSH2 or MLH1 (Fig 1A). In the other four cases the tracings were suggestive of deleted or duplicated exons but other features hinted at artifacts of various types. These features included abnormalities of the control peaks, extra peaks, and background stutter. Repeat experiments failed to confirm the abnormalities in these four cases. These artifacts most likely were due to inferior quality of the templates used initially.

To confirm the presumptive deletions in cases 1-6 and to characterize their breakpoints it was clear that working with diploid DNA would be problematic. In case 1 the breakpoints were likely to lie in the large introns 7 (15.6kb) and 8 (17.3kb). In cases 2-5 the 5’ breakpoint had to be upstream of exon 1, but the extent of the deletions could not be guessed. We therefore assumed that spanning the deletions might not work even with long-range PCR. Preliminary experiments confirmed this. Therefore in each case we obtained haploid DNA by the conversion technique (Yan et al., 2000, Nakagawa et al., 2002), and coarse-mapped the breakpoints by applying several PCRs within introns 7 and 8 (case 1), intron 7 (3’ breakpoint of case 5), intron 6 (3’ breakpoint of cases 3 and 4) and intron 2 (3’ breakpoint of case 2). Once we had determined the 3’ breakpoints of cases 2-5, to search for their 5’ breakpoints we employed inverse PCR technique for haploid DNA. Having thus determined the precise site of each breakpoint we designed deletion-specific PCR using the primer pairs shown in Table 1 to amplify DNA spanning the breakpoints for clinical use. These findings are summarized in Fig 1B and Table 1. The lengths of the deletions were approximately 3 kb in case 1, 11 kb in case 2, and 20 kb in cases 3 and 4 which turned out to be identical. In case 5 the 5’ breakpoint was located approximately 47 kb upstream of exon 1 producing a deletion as long as approximately 84 kb. Interestingly, this deletion comprises all of another gene, tumor-associated calcium signal transducer 1 (TACSTD1). The function of this gene is unknown and in addition to Lynch syndrome this family did not show any specific phenotype that can be explained by hemizygosity for this gene. In all these MSH2 deletions, the breaks appeared to be Alu-mediated. In case 6 MLPA suggested a deletion of exons 3-6 of MLH1 (Fig. 1). The involved introns of MLH1 are much smaller, so long-range PCR between exons 2 and 7 served to identify the breakpoint in diploid as well as haploid DNA. The chosen primers (Table 1) amplified approximately 2.0 kb from the affected allele while the genomic length of this region is approximately 15 kb. Sequence analysis of the abnormal product showed that this deletion spanned 12.6 kb and was mediated by non-homologous recombination.
Figure 1. (A) Screening for deletions in MSH2 and MLH1 using MLPA. In each panel the tracing of the patient was superimposed on that of a control. Arrows indicate presumptive deletions in that the height of the peak from patient DNA is approximately half that of the control. (B) Scheme of five MSH2 genomic deletions and the gene structure upstream to MSH2. Exons of MSH2 are numbered. The sizes of some introns and five genomic deletions are indicated. Alu repeats are shown as thin black bars. More than 50% of the 5’ region of MSH2 (exons 1-9) and the region upstream of MSH2 is comprised of Alu repeats.

DISCUSSION

Our results confirm the previously established fact that large genomic rearrangements of both MSH2 and MLH1 occur with a prevalence that is high enough to make it desirable to look for them in all patients who may have Lynch syndrome. While Southern hybridization would seem to be the obvious choice of screening method, several circumstances argue against it. First, the method requires substantial amounts of DNA, and is work-intensive. Second, in spite of encouraging results (Wijnen et al., 1998) both the sensitivity and specificity of Southern hybridization to detect deletions and other rearrangements are suboptimal. As an example, the 3.5 kb deletion of MLH1 that encompasses exon 16 (Nyström-Lahti et al., 1995) was not detectable even by specially designed Southern hybridization experiments (unpublished observations). Third, once an abnormal Southern blot has been found, it remains to be determined whether it might be due to a polymorphism in a restriction site. If it is due to a deletion, the abnormal band pattern does not necessarily provide a clue to the extent of the deletion (Pyatt et al., 2003). Thus we do not consider Southern hybridization the method of choice to screen for genomic rearrangements of the mismatch repair genes. We also do not favor semiquantitative multiplex exon-by-exon PCR. In our experience, the sensitivity of this method was acceptable when reagents and PCR conditions were highly...
optimized and all phases of the study scrupulously well performed. However, the specificity was relatively low, in
that false positive findings were frequent (data not shown). In striking contrast, MLPA reportedly is highly
sensitive and specific (Schouten et al., 2002; Hogervorst et al., 2003; Montagna et al., 2003). Our present results do
not allow us to critically assess the sensitivity of MLPA to detect deletions and duplications. However, in
preliminary experiments we studied 4 deletions and 2 duplications in the two genes that we had previously
characterized, and MLPA disclosed all 6 without ambiguity. The only way in which the sensitivity can be
objectively determined would be a blind comparison with a “gold standard” technique. In our opinion neither
Southern hybridization nor any other method meets the criteria of a “gold standard” (see above). Nevertheless, our
findings taken together with those from 4 previous studies (Schouten et al., 2002; Gille et al., 2002; Hogervorst et
al., 2003; Montagna et al., 2003) appear to suggest a high sensitivity indeed.

Regarding the specificity of the scanning method, it might be argued that it is of the order of 60% as 4/10
abnormal tracings did not appear to have any deletions. However, this is not a valid interpretation, as the
chromatograms of these cases immediately signaled that the aberrations were likely artifacts, probably due to
inferior quality of the template. For instance, extra peaks occurred, the control peaks were abnormal; there was
background stutter signal and/or exons from both MSH2 and MLH1 appeared abnormal. We anticipate that as
more studies using the MLPA technique will appear, its true sensitivity and specificity will be determined.

Presently, in summary, MLPA appears to be not only simple to perform, but also accurate enough to fulfill
reasonable requirements for a method that is applicable to clinical practice.

We consider it absolutely necessary to determine precisely the breakpoints of all deletions for two main reasons.
First, while the interpretation of MLPA results is straightforward in principle, it is based on dosage estimation
where a deviation from the 1:1 ratio of peak height can be due to other factors than genomic dosage. For instance,
a nucleotide substitution located in the template sequence covered by a probe sequence may interfere with the PCR
reaction and create a false impression of deletion. Another reason for wishing to determine the breakpoints of all
rearrangements is to allow a simple diagnostic PCR to be devised for each rearrangement. This can then be used to
search for the same rearrangement in at-risk family members. We note that breakpoints are highly variable, as
shown by others (Wijnen et al., 1998; Charbonnier et al., 2002) and in this paper. While many deletions of MSH2
encompass exon 1, only a small number of these deletions have been characterized so far, because obviously it is
difficult to clone a distant breakpoint by simple long-range PCR. A case in point is our case 5 in which the deletion
encompasses exons 1-7, but the 5' breakpoint is > 40 kb upstream of exon 1 so that an entire other gene
(TACSTD1) is included in the deletion (Fig 1B). However, once the breakpoints were localized, even in this case a
simple diagnostic PCR reaction could be established.

Increasingly, founder mutations in the mismatch repair genes are being described. Typically founder mutations
occur in isolated populations such as the 3.5 kb exon 16 deletion of MLH1 in Finns (Nyström-Lahti et al., 1995) or
the 1906G>C mutation of MSH2 in Ashkenazi Jews (Foulkes et al., 2002). In such populations, the proportion of
all patients who have the mutation is sometimes high enough (>50%, and ~30%, respectively in the two examples)
to make it useful to screen for the mutation as a first measure in the study of patients with mismatch repair
deficient tumors (Aaltonen et al., 1998). Naturally, such screening methods must be highly sensitive and specific,
such as the PCR we have devised for the five deletions studied by us (Table 1). A case in point is the exon 1-6
deletion we found in cases 3 and 4. The same deletion was found and characterized by another laboratory and
turned out to be present in as many as 9 ostensibly unrelated patients in the United States (Wagner et al., 2003).
It remains to be determined whether it might be common enough to be screened for as a primary clinical test in the
United States.

ACKNOWLEDGMENTS

We thank Dr. R. Fodde for sharing information on one breakpoint, Dr. G. Wiesner and Ms. R. Nagy for
referring a patient, members of the Clinical Cancer Genetics Program at the James Cancer Hospital, the Ohio State
University for help, and Barbara Fersch for preparing this manuscript.

REFERENCES

Chapelle A. 1998. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the


