Original Article

Recurrent and founder mutations in the PMS2 gene


Germline mutations in PMS2 are associated with Lynch syndrome (LS), the most common known cause of hereditary colorectal cancer. Mutation detection in PMS2 has been difficult due to the presence of several pseudogenes, but a custom-designed long-range PCR strategy now allows adequate mutation detection. Many mutations are unique. However, some mutations are observed repeatedly across individuals not known to be related due to the mutation being either recurrent, arising multiple times de novo at hot spots for mutations, or of founder origin, having occurred once in an ancestor. Previously, we observed 36 distinct mutations in a sample of 61 independently ascertained Caucasian probands of mixed European background with PMS2 mutations. Eleven of these mutations were detected in more than one individual not known to be related and of these, six were detected more than twice. These six mutations accounted for 31 (51%) ostensibly unrelated probands. Here, we performed genotyping and haplotype analysis in four mutations observed in multiple probands and found two (c.137G>T and exon 10 deletion) to be founder mutations and one (c.903G>T) a probable founder. One (c.1A>G) could not be evaluated for founder mutation status. We discuss possible explanations for the frequent occurrence of founder mutations in PMS2.

Conflict of interest
Authors declare no conflict of interest.

Mutations in at least four mismatch repair genes cause Lynch syndrome (LS), a condition that predisposes to colorectal and endometrial cancer and to a lesser degree to a number of other cancers (1). Even though the proportion of all colorectal cancer caused by LS is a modest ~3% (but greater if diagnosis is at young age) there is a need for improved strategies to diagnose LS because clinical surveillance and prophylactic surgery can greatly reduce cancer morbidity and mortality (2–4). Different strategies have been devised to detect as many LS mutation carriers as possible as cost-effectively as possible and in many institutions, it is a standard practice to perform immunohistochemistry (IHC) for the mismatch repair proteins as a first step in screening for LS (5–8).

While standard mutation detection methods apply well to MLH1, MSH2, and MSH6, testing for PMS2 gene mutations has been problematic due to the presence of numerous pseudogenes. The use of carefully designed long-range PCR to avoid amplifying the pseudogenes has virtually solved the problem (9–12) so that presently, all four mismatch repair genes can be readily studied for mutations. Deletions in these mismatch repair genes are relatively common; therefore, multiplex ligation-dependent probe amplification is also commonly used. Among the four mismatch repair
genes, mutations in two (MLH1 and MSH2) cause high lifetime risks (penetrance) and together account for some 60–80% of all LS. Mutations in the other two genes (MSH6 and PMS2) have lower penetrance and each accounts for some 10–20% of all LS (13–18).

This communication deals with mutations in the PMS2 gene that were observed multiple times. Mutations that are observed in ostensibly unrelated individuals can be of either recurrent (repeated spontaneous de novo occurrence; also known as ‘hot spot’ mutations) or founder type (inherited from a shared ancestor).

Material and methods

Patients

This study is an extension of a previous study in which 99 probands with colon and/or endometrial cancer who demonstrated isolated absence of tumor staining for PMS2 by IHC were analyzed for PMS2 mutations (16). In total, 61 of the 99 probands (61%) had deleterious mutations (55 monoallelic; 6 biallelic). Of these, 36 were distinct mutations; 25 occurred in just one proband each, 5 occurred in 2 ostensibly unrelated probands and one (c.736_741del6ins11), a previously described ancient founder mutation (10), occurred in 12 ostensibly unrelated probands. This article describes four of the remaining five mutations which respectively occurred in seven, three, three and three ostensibly unrelated probands each (16). For this analysis, we have included an additional seven previously unreported probands with these four PMS2 mutations (total 21 probands with four different mutations). For the fifth mutation, namely the complete gene deletion that occurred in three probands in Senter et al. (16) a DNA sample from only one patient was available; therefore, this mutation was not studied.

Samples from 21 subjects were studied. All research was conducted under approval of the Institutional Review Board (IRB) at the Ohio State University. Fourteen of these subjects were described previously (16) and five of these previously described subjects were accrued anonymously through research collaborations with the Australian Registry of the National Cancer Institute-funded Colon Cancer Family Registry (19). Anonymized samples from an additional seven subjects were provided from the ARUP Laboratories (Salt Lake City, UT) (11). All subjects studied here had LS-associated tumors displaying absence of PMS2 protein with retention of MLH1, MSH2 and MSH6 protein. To determine the population frequency of the mutation-associated haplotypes, we genotyped 80 control individuals. These samples were randomly drawn from Caucasians belonging to a collection of samples obtained from residents of central Ohio for the purpose of serving as controls for genetic studies.

Exon 10 deletion mutant breakpoint analysis

To determine breakpoints for patients with exon 10 deletions, patient DNA was first amplified by long-range PCR using TaKaRa LA Taq and primers specific for PMS2, spanning exon 8 to exon 11. Long-range amplicons were diluted 1:10 and used as template for nested PCR using primers spanning the breakpoint region. Amplicons were then sequenced using BigDye Terminator chemistry on the ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) and compared to NC_000007.13, complement positions 6012870..6048737.

Recurrent and founder mutations

To characterize the haplotypes present in cases and controls, we utilized five out of six microsatellite markers and seven out of nine single nucleotide polymorphisms (SNPs) previously reported (10) that span the PMS2 locus. The available subjects and 80 controls were typed for these PMS2 markers.

In order to prevent the amplification of pseudogenes, DNA samples were amplified using a previously described long-range PCR procedure (9, 11). Amplicons spanning exons 1–5 (long-range amplicon LR1) and 7–9 (LR2) were generated using the previously published primers. For the region encompassing exons 11–15, rather than generating two long-range products spanning exons 10–12 and 12–15, we used the forward primer located in exon 10 and the reverse primer located 3′ of exon 15. This generates an 18,341 bp product (LR3) (11). With this design modification, all the long-range products have at least one primer anchored in an exon not present in any of the pseudogenes.

Using each of the long-range primer sets, 100 ng of DNA were amplified in 25 μl reactions containing 0.2 μM each primer, 1.25 μTaq DNA polymerase (TaKaRa Bio Inc., Otsu, Shiga, Japan), 1× LA PCR Buffer II, and 400 μM each dNTP. Cycling consisted of an initial denaturation at 94°C for 1 min and 30 cycles of 10 s at 98°C and 10 or 18 min (for LR2 or LR3, respectively) at 68°C. Final elongation entailed 10 min at 72°C. The amplification result of long-range PCR was confirmed by gel electrophoresis and diluted (1 in 20) before marker-specific amplification.

Microsatellite markers were typed either by direct labeling of a PCR primer or by utilizing a labeled M13 primer in conjunction with an M13-tailed, ampiclon-specific primer in a three primer PCR. Each 15 μl PCR reaction contained 7.5 μl AmpliTaq Gold master mix (PE Applied Biosystems, Foster City, CA), 100 ng genomic DNA, 10 pmol untailed primer, 5 pmol M13-tailed primer, and 10 pmol FAM-labeled M13 primer. Reactions were cycled using the following profile: 96°C for 10 min, 36 cycles of 96°C for 30 s, 58°C for 30 s, 72°C for 30 s, and final extension at 72°C for 5 min. The PCR product was sized using an ABI 3730 DNA Analyzer (Applied Biosystems).

For the genotyping of SNPs, we used the same PCR conditions as above in the presence of 10 pmol forward and reverse primer with the appropriate long-range PCR product or genomic DNA used as template. The PCR product containing the SNP was subjected to the SNaPshot reaction (Applied Biosystems). The
sequences of the primers used for microsatellite genotyping have been published (10), whereas sequences for the primers used in the SNaPshot reaction can be obtained upon request.

**Haplotype construction**

Genotyping data were used to construct haplotypes using the PHASE 2.0 program (20) according to the manual. Haplotypes associated with the mutation are shown.

**Results**

Genotyping of control samples

We genotyped the controls and used the data to calculate the frequency of the alleles (see Fig. 1). The controls did not have any of the four mutations that we discuss in this study. Using PHASE we constructed haplotypes in the controls. The haplotypes associated with each mutation were searched for in the controls and the number reported.

**c.137G>T**

We had access to samples from a total of 10 mutation-positive probands with c.137G>T, 6 from the original series (16) and an additional 4 from the ARUP collection. The results (Fig. 1) show that a disease-associated haplotype comprising some 375 kb was shared by all subjects. The haplotype stretches from microsatellite D7S481 upstream of exon 1 to SNP rs1468996 downstream of exon 15. This 375 kb shared haplotype was observed in 1 of 80 control individuals from the central Ohio area. All subjects were Caucasian and while...
ancestral information was not available for all probands, three probands reported ancestry in the United Kingdom and another reported Australian ancestry.

Exon 10 deletion

The exon 10 deletion was found in three unrelated probands from Australia (16). For this study, we had access to an additional two probands and a mutation-carrying sister of the third proband, all from the ARUP collection. Breakpoints were confirmed (c.989-296_144+706del) and were identical to breakpoints previously reported in individuals with exon 10 deletions (21, 22). A shared haplotype extending from rs7788441 in intron 7 to microsatellite marker Clen37 downstream of exon 15 was observed in six probands (Fig. 1). The same haplotype occurred in 2 of 80 controls. These data are consistent with a relatively short shared ancestral haplotype. All subjects were Caucasian with unknown ancestral origin.

c.903G>T

The c.903G>T mutation leads to the skipping of exon 8 (16). We studied all three probands and detected a shared haplotype spanning from microsatellite marker Clen35 upstream of exon 1 to SNP rs1468996 some 280 kb downstream of PMS2. This haplotype was seen in 0 of the 80 controls. All subjects were Caucasian and two of three probands reported ancestry from Austria, Hungary, and Germany.

c.1A>G

Of the three probands originally detected carrying this mutation, two were available for study. A shared haplotype was seen from microsatellite marker Clen35 upstream of exon 1 to D7S2201 some 390 kb downstream of PMS2. This haplotype was seen in 7/80 controls. Ancestral information is known for only one of the two subjects, being mixed Irish, French, and Native American.

Discussion

The existence of numerous pseudogenes has made it more difficult to search for mutations in PMS2 than in the other three MMR genes. As a consequence, data on the proportion of all LS that is caused by PMS2 mutations are scarce. Moreover, the documented low penetrance of PMS2 mutations relative to the penetrance of the MMR genes MLH1 and MSH2 (see below) means that PMS2 mutations will be underdiagnosed in the clinical setting where mutation analyses typically are applied to individuals displaying the ‘high risk’ features of strong family history of early onset LS cancers. For these reasons, the proportion of LS caused by mutations in PMS2 can best be estimated by studies in which unselected cases of LS-associated cancers are screened. The data summarized in a large review (7) show MLH1 and MSH2 involvement in 32% and 39%, respectively, while PMS2 and MSH6 were reported to be present in 15% and 14%, respectively, of all LS cases. An additional population-based study of 500 colorectal cancer (CRC) cases (23) disclosed 18 LS probands with a similar distribution of mutations. Thus, as an overall conclusion MSH2 and MLH1 account for ~70% of diagnosed LS while MSH6 and PMS2 together account for the remaining ~30% (7). From a practical point of view, these numbers suggest that to adequately assess the presence of LS all four genes must be considered.

This study focuses on those mutations in PMS2 that occurred repeatedly in a series of 99 probands whose tumors did not stain for PMS2 protein by IHC (16). The subjects emanated from numerous institutions mainly in Northern Europe, North America and Australia, being mostly Caucasians of European origin. It is important to bear in mind that we cannot, therefore, make inference about other ethnicities or nationalities. Moreover, because the initial series of 99 ostensibly unrelated probands contained at least 24 population-based probands while at least 19 probands were from high-risk clinics (exact numbers are not available), there may be a bias in favor of higher rather than lower penetration mutations if such exist. Nevertheless, with these limitations our series of subjects is by far the largest of its kind and therefore allows at least some tentative conclusions of population relevance.

We show that repeated mutations in PMS2 are common and whenever feasible to assess, are likely to be of founder nature. Among the 61 probands, 31 carried a mutation seen in at least three probands and one mutation was observed in seven probands. In addition, as shown in Table 1 several of these mutations have been seen and published in patients who were not part of the initial series described in Senter et al. (16). Thus, it appears that approximately half of all PMS2 mutations occur repeatedly in the Caucasian population.

The most common mutation described in Senter et al. (16), c.736_741del6ins11, has been studied in detail previously (10). The second most common mutation described in Senter et al. (16), c.137G>T (Ser46Ile) has the characteristics of a deleterious missense change (24, 25). Our data allow us to conclude that this mutation is inherited from a single shared ancestor. The haplotype is short suggesting that the mutation occurred many generations ago, but with the limited number of affected individuals available for study we are not able to estimate the age of the mutation with any degree of precision. Additionally, the exon 10 deletion (the second mutation examined in our study) very likely is of a founder type, due to the presence of shared breakpoints and a shared haplotype. For this shared haplotype, we were unable to determine the upstream start. As this haplotype was never detected.

Recurrent and founder mutations
in controls and although we have studied only three probands we conclude that this is a probable founder mutation.

The last mutation analyzed (c.1A>G) disrupts the first translation initiation codon leading to 5' truncation of the putative protein. Only two samples were available from subjects with this mutation and they shared an even longer haplotype which was present in 7/80 controls. In this case too, we were unable to determine the upstream start of the shared haplotype. Thus, while shared ancestry is a distinct possibility, this mutation could also be the one that recurs frequently de novo. We note here that de novo mutations are rare in the mismatch repair genes (26).

Founder mutations are not unique to PMS2. Founder mutations are well known in MLH1 (27–31) and MSH2 (32–34). At least one recurrent ‘hot spot’ mutation is widespread worldwide. This is the intronic MSH2 c.942+3A>T splice site mutation that apparently arises frequently de novo as a result of meiotic misalignment at a stretch of 26 adenosines in the 5’ region of intron 5 (35). Are founder or recurrent mutations less common in MLH1, MSH2 and MSH6 than in PMS2? The large multicenter study on MSH6 by Baglietto et al. (18) identified 74 distinct mutations in a total of 113 probands. Among the 74 mutations, 22 were observed in more than one proband (range = 2–6 probands). In total, 29/113 families displayed mutations seen more than twice, as compared to the 31/61 noted by us for PMS2. Thus, it is possible that the two MMR genes with the lowest penetrance (PMS2 and MSH6) also share the property of having frequent recurrent or founder mutations, but they may be more abundant in PMS2 than MSH6. It is documented that the penetrance of cancer is lower in PMS2 (lifetime risk of CRC ∼20%) than in MLH1 and MSH2 (lifetime risk of CRC ∼40–60%) [Senter et al. (16) and references cited within]. Unfortunately, data establishing the proportion of repeated mutations in MLH1 and MSH2 are not readily available. We are not aware of publications in which the occurrence of mutations has been determined in large numbers of probands from panmixing (as opposed to isolated; geographically distinct) populations. Nevertheless, population-based studies reviewed in Palomaki et al. (7) list the MMR mutations found in altogether only 82 probands with LS. These data are too few to conclude anything with certainty about the proportion of founder mutations in the two most prevalent MMR genes compared to PMS2 and MSH6. We suggest that further, much larger population-based studies are desirable to shed light on this question.

Founder mutations are believed to become enriched by at least two alternative mechanisms. First, if a rare mutation is introduced into an isolated population that subsequently expands without significant influx of genes, it can become enriched simply by genetic drift. (More often, however, it can decrease or become extinct from genetic drift.) This mechanism is believed to account for those numerous examples of frequent founder mutations seen in Icelanders, Finns, Ashkenazi Jews, French Canadians, and other typical isolated founder populations. This mechanism does not readily apply to our findings in PMS2 which are derived from large panmixing Caucasian populations. Another well-known cause by which a particular mutation can become enriched occurs when its effect carries an advantage (positive selection). This mechanism is well known for example from the hemoglobin gene where heterozygosity for the most common sickle cell anemia mutation confers protection against malaria (36). We are unaware of any evidence about positive selection of mutations in PMS2.

In summary, founder mutations appear to be common in PMS2. As more PMS2 mutations are identified through population-based screening of colon and/or endometrial cancers using IHC followed by appropriate germline genetic testing, more PMS2 mutation carriers are likely to be identified and could provide much more detailed estimates of the prevalence of these mutations. It is possible that if certain mutations are identified in a significant number of patients, standard methodology of PMS2 mutation detection could be altered by testing for common mutations before sequencing the entire gene.

Acknowledgements

This work was supported by the National Cancer Institute (grant number: CA16058 and CA67941), National Institutes of Health under RFA # CA-95-011 and through cooperative agreements with members of the Colon Cancer Family Registry (Colon CFR) and P.I.s. Samples and corresponding data were provided by Australasian Colorectal Cancer Family Registry (U01 CA097735). We thank Shuying Sun, Jennifer Panescu, James Mackay, Joy Larsen Haidle, Marc Greenblatt, Daniel Buchanan, Graham Giles, Graham Suthers, Graham Young and Jack Goldblatt for help and Paul Fuerst for helpful discussion.

**Table 1. PMS2 mutations found in two or more of 61 probands studied by Senter et al. (16) and number of probands with the same mutations reported in the literature**

<table>
<thead>
<tr>
<th>Mutation</th>
<th># Probands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Senter et al. (16)</td>
</tr>
<tr>
<td>c.736_741del6ins11 (P246CfsX3)</td>
<td>12 (11, 37)</td>
</tr>
<tr>
<td>c.137G&gt;T (S46I)</td>
<td>7 (11, 37–40)</td>
</tr>
<tr>
<td>Deletion exon 10</td>
<td>3 (11, 21, 22)</td>
</tr>
<tr>
<td>(c.989–296_1144+706del)</td>
<td>3</td>
</tr>
<tr>
<td>c.903G&gt;A (E705K)</td>
<td>2</td>
</tr>
<tr>
<td>c.1831_1832insA</td>
<td>2 (11)</td>
</tr>
<tr>
<td>(I611NfsX2)</td>
<td>–</td>
</tr>
<tr>
<td>c.2113G&gt;A (E705K)</td>
<td>2</td>
</tr>
<tr>
<td>c.949C&gt;T (Q317X)</td>
<td>2</td>
</tr>
<tr>
<td>Deletion of exons 5, 6, 7</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note**
- This mutation occurred monoallelic in one proband and homozygous biallelic in one proband. In the latter case, the parents were first cousins and each was heterozygous for the mutation.
References


Recurrent and founder mutations