Genetic linkage analysis in hereditary non-polyposis colon cancer syndrome

Nicola J Froggatt, Jane Koch, Rhodri Davies, D Gareth R Evans, Andrew Clamp, Oliver W J Quarrell, Jean Weissenbach, Shirley V Hodgson, Bruce A J Ponder, David E Barton, Eamonn R Maher

Abstract
Hereditary Non-polyposis Colon Cancer Syndrome (HNPCC) is the most common cause of familial colorectal cancer. Molecular genetic studies of HNPCC have shown evidence of locus heterogeneity, and mutations in four genes (hMLH1, hMSH2, hPMS1, and hPMS2) which encode components of the mismatch enzyme repair system may cause HNPCC. To determine the extent and nature of locus heterogeneity in HNPCC, we performed genetic linkage studies in 14 HNPCC families from eastern and north-western England. Linkage to hMLH1 was excluded in six families, each of which were likely to be linked to hMSH2 (lod score >1.0 in each family and total lod score for all six families = 7.64). Linkage to hMSH2 was excluded in three families, each of which were likely to be linked to hMLH1 (lod score >1.9 in each family and total lod score at hMLH1 for all three families = 3.93). In the remaining five families linkage to hMSH2 or hMLH1 could not be excluded. These results confirm locus heterogeneity in HNPCC and suggest that, in the population studied, most large families with HNPCC will have mutations in hMSH2 or hMLH1. We did not detect any correlation between clinical phenotype and the genetic linkage results, but a Muir-Torre syndrome family excluded from linkage to hMLH1 was likely to be linked to hMSH2 and showed microsatellite instability in a tumour from an affected relative.


It is estimated that up to 10% of all cases of colorectal cancer occur in genetically predisposed people. The importance of inherited susceptibility is greater in early onset cases such that people with a first degree relative with colorectal cancer diagnosed <45 years have a five-fold increase in lifetime risk of colorectal cancer. A small proportion of familial colorectal cancer is caused by familial adenomatous polyposis coli (FAPC), a dominantly inherited disorder with an incidence of 1 in 8000 caused by mutations in the APC gene on chromosome 5q21. FAPC is characterised by the development of numerous (>100) colonic polyps in the second and third decades, and colorectal carcinoma is almost inevitable unless colectomy is performed. A much larger proportion (up to 10%) of cases are thought to be caused by hereditary non-polyposis colon cancer syndrome (HNPCC). This disorder is characterised by a dominantly inherited predisposition to early onset colon cancer (mean 46 years). Affected patients develop multiple colorectal cancers, but florid colonic polyposis is not a feature. In many cases there is also an increased risk of extracolonic cancers such as uterine, ovarian, and ureteric. Clinical heterogeneity is well recognised in HNPCC, and families have been subdivided according to the presence or absence of extracolonic cancers (Lynch family cancer syndromes types 2 and 1 respectively), or associated skin lesions (Muir-Torre syndrome). Families with a classical HNPCC phenotype do not show linkage to the APC gene, but Peltomaki et al mapped a HNPCC locus to chromosome 2p16 in 1993. In addition, these investigators also detected evidence of widespread instability of simple DNA repeat polymorphisms in tumours from HNPCC patients suggesting that HNPCC can result from defects in the mechanisms responsible for maintaining the fidelity of DNA replication. Subsequent studies identified the chromosome 2p16 HNPCC gene (hMSH2) as the human homologue of the bacterial mutS. Further studies have shown that mutations in three other genes (hMLH1 in chromosome 3p21, hPMS1 in chromosome 2q31-32, and hPMS2 in 7p22) which encode components of the mismatch repair enzyme complex may also produce HNPCC. Despite the rapid progress in the identification of HNPCC genes, the molecular genetics of HNPCC have been investigated in relatively few families. The initial study which successfully mapped a HNPCC locus to 2p16 included 14 kindreds from Europe, North America, and New Zealand and it was estimated that approximately 50% of HNPCC families were linked to hMSH2. However, most Finnish families are not linked to hMSH2 and the majority appear to be caused by mutations in hMLH1. Clearly the optimum approach for the molecular genetic diagnosis of HNPCC will depend on the relative importance of the four candidate genes in the relevant population. We have investigated the nature and extent of locus heterogeneity in HNPCC by performing genetic linkage studies in 14 families ascertained from eastern and north western England.
Materials and methods

PATIENTS

One hundred and sixty eight subjects (45 affected patients and 123 at risk relatives and spouses) from 14 families with HNPCC were genotyped with microsatellite markers linked to hMSH2 (D2S119, D2S391, D2S288, CA5, D2S123, D2S378, D2S136) and hMLH1 (D3S1029, D3S1298, D3S161, and D3S1619). Families were ascertained from East Anglia and north west England and were of British or Irish origin. All families satisfied the Amsterdam criteria for the diagnosis of HNPCc: at least three relatives with colorectal cancer in two or more generations, with at least one patient diagnosed aged <50 years. Non-colonic cancers known to be associated with the Lynch 2 phenotype (uterus, ovary, and ureter) occurred in seven families. A diagnosis of Muir-Torre syndrome was made in one kindred (C004). No families had evidence of classical or attenuated adenomatous polyposis coli. Family members were considered affected if they developed (1) colorectal cancer, (2) an extraintestinal cancer typical of HNPCc (uterine or ovarian), or (3) multiple (2 to 20) adenomatous polyps aged <40 years.

DNA ANALYSIS

High molecular weight DNA was isolated from peripheral blood or lymphoblastoid cell lines by conventional methods. DNA was amplified by the polymerase chain reaction under the appropriate conditions and the alleles resolved on a 6% polyacrylamide/6 mol/l urea sequencing gel as described previously.26 Evidence of microsatellite instability was sought in tumours from affected members of five families using five microsatellite polymorphisms: D3S1007, D3S1076, D5S346, D17S588, and CA5 on chromosome 2p.14 DNA was extracted from 15 μm sections from paraffin embedded normal and tumour tissue, according to the following procedure, adapted from Smith et al.: one section was placed into 300 μl extraction buffer (50 mmol/l KCl, 10 mmol/l Tris-Cl pH 8.3, 2.5 mmol/l MgCl2, 0.1 mg/ml gelatin, 0.45% Nonidet P40, 0.45% Tween 20), 200 μg Pronase18 (Boehringer Mannheim UK, Lewes, UK) added, and the extraction incubated at 40°C for seven days. Chelex™ resin (BioRad, Hemel Hempstead, UK) was added to a final concentration of 1-67% from a well mixed 5% stock solution, and the extraction incubated at 56°C for 30 minutes. After enzyme inactivation by boiling for 10 minutes, the extraction was microcentrifuged for 30 seconds at 13 000 rpm and 1 to 5 μl was used for each amplification. The finding of microsatellite instability at two or more loci was considered significant.31

GENETIC LINKAGE ANALYSIS

LI PED and LINKAGE computer programmes were used for two point and multipoint linkage analysis. Age dependent penetrance values were as follows: age <30 years = 0.00, 30–45 years = 0.15, 46–60 years = 0.70, >60 years = 0.90. The risk of an at risk person coincidentally developing sporadic cancer of the HNPCc type was estimated at 0.01 at age <45 years, 0.03 aged 45–60 years, 0.05 aged >60 years, and 0.001 for the occurrence of multiple primary tumours (EUROFAP HNPCc collaboration, personal communication, Professor T Bishop). The risk of multiple adenomatous colonic polyps <40 years in a non-genic carrier was taken as 0.03. Each microsatellite marker was analysed assuming five alleles of equal frequencies. The pre-existing map order and distances (recombination fractions) were derived from published studies14–18 and is shown in fig 1. For each family a haplotype was constructed for the chromosome 2p16 and 3p21 markers analysed and lod scores at hMSH2 and hMLH1 were calculated using the MLINK program and results from informative flanking markers.

Table 1 Two point lod scores at loci on chromosome 2p and 3p for 14 HNPCc families

<table>
<thead>
<tr>
<th>Recombination fraction</th>
<th>0.00</th>
<th>0.01</th>
<th>0.05</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2S119</td>
<td>10.97</td>
<td>7.25</td>
<td>3.77</td>
<td>-0.02</td>
<td>-0.52</td>
<td>-0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>D2S391</td>
<td>-7.79</td>
<td>-5.18</td>
<td>-1.43</td>
<td>-0.26</td>
<td>1.38</td>
<td>-0.09</td>
<td>0.31</td>
</tr>
<tr>
<td>D2S288</td>
<td>-1.79</td>
<td>-1.45</td>
<td>-0.79</td>
<td>-0.41</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>CA5</td>
<td>-5.25</td>
<td>-3.99</td>
<td>-2.15</td>
<td>-1.20</td>
<td>-0.43</td>
<td>-0.17</td>
<td>-0.08</td>
</tr>
<tr>
<td>D2S123</td>
<td>-2.28</td>
<td>-2.53</td>
<td>-2.75</td>
<td>-2.63</td>
<td>-1.96</td>
<td>1.09</td>
<td>0.31</td>
</tr>
<tr>
<td>D2S378</td>
<td>-3.16</td>
<td>-1.12</td>
<td>-1.38</td>
<td>-2.22</td>
<td>-2.17</td>
<td>1.56</td>
<td>0.45</td>
</tr>
<tr>
<td>D2S136</td>
<td>-3.24</td>
<td>-1.91</td>
<td>-0.15</td>
<td>-0.64</td>
<td>-0.98</td>
<td>0.71</td>
<td>0.26</td>
</tr>
<tr>
<td>D3S1029</td>
<td>-1.28</td>
<td>-1.75</td>
<td>-1.57</td>
<td>-0.02</td>
<td>0.28</td>
<td>0.23</td>
<td>0.09</td>
</tr>
<tr>
<td>D3S1298</td>
<td>-1.63</td>
<td>-0.47</td>
<td>-0.50</td>
<td>-0.65</td>
<td>-0.13</td>
<td>-0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Figure 1 Genetic maps of chromosome 2p16 (A) and 3p21 (B) showing the positions of hMSH2 and hMLH1 and marker loci investigated.
Results
GENETIC LINKAGE ANALYSIS
Significant evidence of linkage (total lod score >3·0 for all 14 families) was not detected at any of the 11 loci investigated (table 1). However, inspection of marker haplotypes and of individual family lod scores was consistent with locus heterogeneity. Formal heterogeneity analysis provided clear evidence for locus heterogeneity with linkage to hMSH2 (evidence that some families linked to hMSH2, \( \chi^2 = 32·1, p<2 \times 10^{-5} \)) and hMLH1 (evidence that some families linked to hMLH1, \( \chi^2 = 21·9, p<3 \times 10^{-6} \)). On the basis of these results the families were divided into three groups (table 2).

Group A. Exclusion of linkage to hMLH1 and probable linkage to hMSH2
Linkage to hMLH1 was excluded in six families (C001, C002, C003, C004, C015, C016). In each of these haplotype analysis was consistent with linkage to hMSH2 and the total lod score at hMSH2 for all six families was 7·64 at \( \theta = 0 \) (fig 2). However, although each of these families was likely to be linked to hMSH2, no single family provided a lod score >3.

Group B. Exclusion of linkage to hMSH2 and probable linkage to hMLH1
Linkage to hMSH2 could be excluded in three families (C009, C012, C014), and in each of these families haplotype analysis was consistent with linkage to hMLH1. The total lod score at hMLH1 for these three families was 3·93 although no single family gave a lod score >3. Nevertheless each of the families excluded from linkage to hMSH2 was likely to be linked to hMLH1 (fig 3).

Group C. Not excluded from linkage to hMSH2 or hMLH1
Five families could not be excluded from linkage to hMSH2 or hMLH1 and did not give convincing evidence of linkage to hMSH2 or
hMLH1. There was no evidence for a third HNPCC locus both by formal heterogeneity testing and by direct inspection of hMSH2 and hMLH1 haplotypes.

CLINICAL PHENOTYPE AND GENETIC LINKAGE ANALYSIS
There was no apparent correlation between the results of genetic linkage analysis and clinical phenotype. A Lynch 1 phenotype (no extra-colonic cancers) was found in two families in group A, two families in group B, and two families in group C. One family (C004) in group A had a Muir-Torre syndrome phenotype, and a Lynch 2 phenotype (presence of uterine, ovarian, or ureteric cancer) occurred in three families in group A, one in group B, and three in group C. The mean age (SD) of colorectal cancer in affected subjects was similar in each of three groups: group A, 46.3 (SD 10.3) years; group B, 44.3 (SD 8.8) years; and group C, 44.8 (SD 6.6) years.

REPLICATION ERRORS IN COLORECTAL CANCERS
Replication errors were detected at two or more loci in colorectal cancers from the five families.

**Figure 3**  Pedigree details, cancer diagnosis, and examples of haplotyping construction for loci linked to hMSH2 (top panel) and hMLH1 (lower panel) in family C012. Results were obtained for unaffected subjects marked *, but are not given to preserve confidentiality. Under each symbol is listed the confirmed diagnosis (Co = colorectal cancer, Ut = uterine cancer, Ov = ovarian cancer) and age at diagnosis.
Table 2  Multipoint lod scores at hMSH2 and hMLH1 in 14 HNPCC families (see text for details of groups). The posterior probabilities of linkage to hMSH2 or hMLH1 were calculated according to the best fitting model (54% families resulting from hMSH2 and 46% from hMLH1).

<table>
<thead>
<tr>
<th>Group</th>
<th>Family</th>
<th>hMSH2</th>
<th>hMLH1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lod score</td>
<td>Posterior probability of linkage</td>
</tr>
<tr>
<td>A</td>
<td>C001</td>
<td>1-3</td>
<td>1.00</td>
</tr>
<tr>
<td>A</td>
<td>C002</td>
<td>1-16</td>
<td>1.00</td>
</tr>
<tr>
<td>A</td>
<td>C004</td>
<td>1-11</td>
<td>1.00</td>
</tr>
<tr>
<td>A</td>
<td>C003</td>
<td>1-54</td>
<td>1.00</td>
</tr>
<tr>
<td>A</td>
<td>C015</td>
<td>1-01</td>
<td>1.00</td>
</tr>
<tr>
<td>A</td>
<td>C012</td>
<td>-3.3</td>
<td>0.00</td>
</tr>
<tr>
<td>B</td>
<td>C014</td>
<td>-3.46</td>
<td>0.00</td>
</tr>
<tr>
<td>B</td>
<td>C009</td>
<td>-3.39</td>
<td>0.00</td>
</tr>
<tr>
<td>C</td>
<td>C006</td>
<td>-0.01</td>
<td>0.45</td>
</tr>
<tr>
<td>C</td>
<td>C008</td>
<td>-0.37</td>
<td>0.80</td>
</tr>
<tr>
<td>C</td>
<td>C007</td>
<td>-0.18</td>
<td>0.08</td>
</tr>
<tr>
<td>C</td>
<td>C010</td>
<td>-0.87</td>
<td>0.01</td>
</tr>
<tr>
<td>C</td>
<td>C011</td>
<td>-1.36</td>
<td>-1.00</td>
</tr>
</tbody>
</table>

available for analysis. Four families belonged to group C (C007, C008, C010, C011) suggesting that susceptibility to colorectal cancer in these families was caused by a mutation in a mismatch repair gene. In addition, microsatellite instability was detected in a colorectal cancer from a family (C004) with Muir-Torre syndrome.

Discussion

Although no single family gave a lod score of >3 at hMSH2 or hMLH1, six families were likely to be linked to hMSH2 and three families to hMLH1. This interpretation is supported by exclusion of linkage to hMLH1 in those classified as probably being linked to hMSH2 and exclusion of linkage to hMSH2 in families with lod scores >1 at hMLH1. Our data would suggest that in the population studied most HNPCC families will have mutations in hMSH2 or hMLH1 and that mutations in other candidate HNPCC genes probably account for a minority of cases. Although further data are needed to substantiate these preliminary conclusions it is clear that there are geographical variations in the genetic epidemiology of HNPCC. In addition to the high frequency of hMLH1 families in Finland,16 18 Lindblom et al22 reported linkage to hMLH1 in two Swedish kindreds and excluded linkage to chromosome 2p21 and 3p21 (that is, hMSH2 and hMLH1 respectively) in a third kindred. In a study of seven Canadian families, Green et al35 detected linkage to hMSH2 in two, excluded linkage to hMSH2 in three, and obtained indeterminate results in the remaining two families. We did not identify any families in which linkage to hMSH2 and hMLH1 could be excluded suggesting that germline mutations in hPMS1 and hPMS2 will probably account for a minority of HNPCC cases in the population studied. We note that a recent linkage study comprising 13 large HNPCC families from North America, Finland, and New Zealand reached a similar conclusion.34 The high incidence of hMLH1 linked families in Finland appears to result from a founder effect; however, we did not find a shared haplotype among our hMSH2 or hMLH1 linked families.

The ability to identify the most likely candidate HNPCC gene for each family would facilitate the detection of germline mutations. Although HNPCC families have been divided into Lynch 1 and Lynch 2 subgroups, the classification of individual families is frequently not clear cut and does not appear to correlate with the results of molecular genetic analysis. Nevertheless further research is required to define the molecular basis of phenotypic variation in HNPCC. Although Muir-Torre syndrome is sometimes distinguished from a Lynch 2 phenotype, we and Honchel et al37 also found microsatellite instability in tumours from Muir-Torre kindreds. In addition, Hall et al35 found linkage to hMSH2 in two families with Muir-Torre syndrome, suggesting that the two disorders can be allelic. It is of interest that the Muir-Torre family we investigated was also linked to hMSH2, but Green et al35 excluded linkage to hMSH2 in a Canadian Muir-Torre kindred suggesting that this phenotype is not exclusively caused by mutations in hMSH2.

Although family linkage studies can provide information on the genetic epidemiology of HNPCC, an unsuitable family structure or unavailability of DNA from relatives who have died limits the number of families that can be investigated by genetic linkage analysis. In the presence of marked locus heterogeneity, few families are extensive enough to allow genetic linkage studies to be used for accurate presymptomatic diagnosis. Furthermore the presence of phenocopies (relatives with sporadic colorectal cancer) may further complicate the interpretation of genetic linkage studies, although investigation for microsatellite instability may help to distinguish HNPCC from sporadic colorectal tumours.31 Consequently the clinical application of molecular genetic diagnosis of HNPCC is likely to be based on direct mutation detection. The most efficient strategy will depend on the results of further molecular studies in the relevant population.

We thank the Cancer Research Campaign (CRC) for financial support and Dr Bert Vogelstein for details of CA5. We are indebted to Professor Tim Bishop for performing formal heterogeneity analysis. We are grateful to the many colleagues who have provided us with access to families, and to the patients and relatives who took part in this study.
Genetic linkage analysis in hereditary non-polyposis colon cancer syndrome


