Explaining the Familial Colorectal Cancer Risk Associated with Mismatch Repair (MMR)-Deficient and MMR-Stable Tumors
Lauri Aaltonen,1 Louise Johns,3 Heikki Järvinen,2 Jukka-Pekka Mecklin,4 and Richard Houlston3

Abstract

Purpose: There is a paucity of data quantifying the familial risk of colorectal cancer associated with mismatch repair (MMR)-deficient and MMR-stable tumors. To address this, we analyzed a population-based series of 1,042 colorectal cancer probands with verified family histories.

Experimental Design: Constitutional DNA from probands was systematically screened for MYH variants and those with cancers displaying microsatellite instability (MSI) for germ-line MMR mutations; diagnoses of familial adenomatous polyposis and juvenile polyposis were established based on clinical phenotype and mutational analysis. Familial colorectal cancer risks were enumerated from age-, sex-, and calendar-specific population incidence rates. Segregation analysis was conducted to derive a model of the residual familial aggregation of colorectal cancer.

Results: Germ-line predisposition to colorectal cancer was identified in 37 probands [3.4%; 95% confidence interval (95% CI), 2.4-4.6]: 29 with MLH1/MSH2 mutations, 2 with familial adenomatous polyposis, 1 with juvenile polyposis, and 5 with biallelic MYH variants. The risk of colorectal cancer in first-degree relatives of probands with MSI and MMR-stable cancers was increased 5.01-fold (95% CI, 3.73-6.59) and 1.31-fold (95% CI, 1.07-1.59), respectively. MSH2/MLH1 mutations were responsible for 50% of the overall excess familial risk and 80% of the risk associated with MSI cancers but 32% of the familial risk was unaccounted for by known loci. Genetic models based on major gene loci did not provide a better explanation of the residual familial aggregation than a simple polygenic model.

Conclusions: The information from our analyses should be useful in quantifying familial risks in clinical practice and in the design of studies to identify novel disease alleles.

Family history is acknowledged to be one of the strongest risk factors for the development of colorectal cancer. Numerous studies have documented that ~15% of individuals with colorectal cancer report a close relative also affected by the disease (1–3). A meta-analysis of epidemiologic studies we conducted showed that having a first-degree relative with colorectal cancer approximately doubled the risk of colorectal disease (1–3). A meta-analysis of epidemiologic studies we conducted showed that having a first-degree relative with colorectal cancer approximately doubled the risk of colorectal cancer; and if the affected relative was diagnosed before the age of 45 years, the risk was increased ~4-fold (4). Germ-line mutations in APC, SMAD4, ALK3, STK11/LKB1, MYH, and the mismatch repair (MMR) genes have all been shown to predispose to syndromic forms of colorectal cancer: familial adenomatous polyposis (FAP; Mendelian inheritance in man 175100),2 juvenile polyposis (Mendelian inheritance in man 174900), Peutz-Jeghers syndromes (Mendelian inheritance in man 175200), recessive polyposis (MYH; Mendelian inheritance in man 608456), and hereditary non-polyposis colorectal cancer (HNPCC; Mendelian inheritance in man 120435-6), respectively (5). Aside from the substantive risk associated with HNPCC (6, 7), germ-line mutations in the MMR genes are thought to contribute significantly to the overall burden of colorectal cancer (8). Human MMR is mediated by several genes, including MLH1, MSH2, MSH6, PMS2, MLH3, and MSH3 (9). Together, these MMR genes are responsible for recognizing and initiating repair of mismatched base pairs. Failure of MMR processing has the potential to allow the incorporation of unrecognized base pair mismatches into the genome (10, 11), which may be sufficient to deleteriously affect the expression of genes essential to normal cellular function. Microsatellite instability (MSI) caused by expansion or contraction of short nucleotide repeats is a characteristic of MMR deficiency (11, 12) and is detectable in the majority of colorectal cancers arising in carriers of germ-line MMR mutations (13, 14).
Mutations in two of the MMR genes, MSH2 and MLH1, segregate with disease in ~50% of multigenerational colorectal cancer families, fulfilling the Amsterdam criteria for a diagnosis of HNPPC (14–16). Few instances of germ-line mutations in MSH2 or MLH1 have to date been documented outside the context of HNPPC families, except in individuals diagnosed with colorectal cancer at a young age (17), and only a minority of HNPPC families seems to segregate mutations in other MMR genes, such as MSH6 (18).

Deficiency of MMR also plays a role in the development of colorectal cancer outside the context of HNPPC. Approximately 12% of all colorectal cancers (13), especially those developing in the proximal colon (19), exhibit MSI. The significance of MMR deficiency in such cancers in terms of familial risk is unclear. Some of the cases represent unrecognized HNPPC; however, in the majority of these sporadic colorectal cancers, MMR deficiency is an acquired event, a consequence of somatic mutation or abnormal DNA methylation (20, 21). Whether MMR deficiency in such cancers in terms of familial risk is a consequence of stochastic mechanisms or whether some individuals are genetically susceptible to such processes is unknown.

A recent twin study suggests that ~35% of all colorectal cancer involves an inherited susceptibility (22). Understanding the genetic basis of this inherited susceptibility is relevant to being able to discriminate between high- and low-risk groups and is important in the design of studies to identify novel predisposition loci. Although the cancer risks associated with the major Mendelian colorectal cancer predisposition syndromes are well established, less is known about the relative contribution of the known loci to the overall familial risk of the disease and, most importantly, the magnitude of the residual familial risk. Specifically, there are limited data on the genetic basis of the excess familial risk associated with MMR-competent and MMR-deficient cancers.

To derive unbiased estimates of familial colorectal cancer risks stratified by MMR status and determine the contribution of known genes to the overall disease burden, we analyzed a large population-based series of colorectal cancer probands with verified family histories.

### Materials and Methods

**Subjects and pedigrees.** Family histories of first-degree relatives (parents, siblings, and offspring) were obtained from 1,042 consecutive colorectal cancer probands ascertained from nine large regional hospitals in southeastern Finland. Verification of all reported cancers in families was sought through reference to cancer registry data, death certificates, and medical records. Further details about the design and conduct of the study are described in previously published material (8, 23, 24). Follow-up for the entire cohort is complete to the end of 1997. Colorectal cancer was defined according to the ninth revision of the International Classification of Diseases by codes 153 to 154 (25).

Informed consent for the study was obtained from participants, and the study was carried out with ethical review board approval in accordance with the tenets of the declaration of Helsinki.

**Molecular analyses.** The MMR status of cancers from probands and the presence of MYH variants were determined in previous studies (23, 24, 26). MMR deficiency in cancers was detected by screening tumors for MSL. Briefly, 236 of the tumors were genotyped using 7 to 14 radioactively labeled markers, the subsequent 273 tumors were genotyped with 7 to 16 fluorescently labeled markers, and the remainder were genotyped with the two poly(A) markers BAT26 and TGFBI. Under the scheme proposed by the National Cancer Institute for classifying instability in colorectal cancer (27), all tumors referred to as displaying MSI in this report are compatible with a diagnosis of MSI high. Probands with MSI tumors were sequenced for MLH1 and MSH2 germ-line mutations. Mutational analysis of MSH6 was not systematically undertaken.

Identification of the MYH sequence variants Y165D and G382D in the 1,042 probands was undertaken adopting a DNA pooling strategy and solid-phase minisequencing (26). Probands with one mutated allele in the germ-line were sequenced for all exons to search for additional pathogenic sequence changes. The diagnosis of FAP and juvenile polyposis in probands was made based on clinical phenotype supported by mutational analysis of the relevant gene.

**Statistical analyses.** In all analyses, a P value of ≤0.05 was considered to be statistically significant. Confidence limits for the prevalence of MSI in cancers and germ-line mutations were computed under the assumption that frequencies followed a Poisson distribution. The association between categorical variables was made using either Fisher’s exact test or χ² test. The Armitage statistic was used to test for trends in categorical variables. Differences in randomly distributed variables were assessed by means of either Mann-Whitney U test or t test.

The cumulative probability of colorectal cancer in defined first-degree relatives of probands, \( P(\tau | t) = 1 - S(\tau) \), where \( S(\tau) \) is the survivor function, was estimated by the Kaplan-Meier product-limit method (28). Relatives were censored at the date of pedigree ascertainment, emigration, or last contact with the proband. Cancer incidence in relatives was truncated at age 84 years to avoid the inherent problems of misclassification of registered cancer diagnosis and death, which, if discounted, represents a source of bias. The variance of the product-limit estimator was estimated according to the formulae of Greenwood (29) and was used to calculate confidence intervals. Differences in colorectal cancer risk according to age at diagnosis of colorectal cancer in proband, sex and type of relative, and MSI status were assessed by means of the log-rank test. We further tested for risk differences by using Cox’s proportional hazards regression model (30) to adjust for potential demographic confounding factors, such as birth year and sex. We used a robust variance procedure (31) to adjust for intrafamilial correlations. Comparison of the risk of cancer in relatives with the risk in the general population was based on rates abstracted from the Finnish Cancer Registry using the following formula: \[ 1 - \exp[-\lambda \cdot t] \] ref. 32. These analyses were conducted using the statistical software program Stata version 8 (Stata Corp., College Station, TX).

Expected numbers of colorectal cancer cases in first-degree relatives of probands before age 85 years were computed using age, sex, and calendar period incidence rates for the general population of Finland using the person-years program (33). Observed numbers were compared with numbers by means of the standardized morbidity ratio (SMR) assuming a Poisson distribution using two-sided P values to assess significance.

To model the genetic basis of the residual familial aggregation of colorectal cancer, segregation analysis under a mixed model (34) was conducted using the pedigree analysis program (35). Analyses were conducted using the disease status, defined as a dichotomous trait, affected or unaffected with colorectal cancer, and age at diagnosis or at last observation. To adjust for ascertainment, the likelihood for each pedigree was conditioned on the proband being affected at age at diagnosis of colorectal cancer. Liability classes were defined by population risks of colorectal cancer. Models were compared using the likelihood ratio test and Akaike’s information criterion (36) defined by the following: \( 2 \times (\log \text{maximum likelihood} + \text{number of variables estimated}) \) The Akaike’s information criterion assessing the relative fits of models was used by adding a penalty to each log likelihood to reflect
the number of variables estimated under the specific model. The most parsimonious model was taken to be that generating the minimum Akaike's information criterion value.

**Results**

**Descriptive statistics.** Five hundred and twenty-seven (50.4%) of the 1,042 probands were male and 517 (49.6%) were female; 613 (58.8%) had been diagnosed with colon cancer (ninthy revision of the International Classification of Diseases code 153) and 429 (41.2%) with rectal cancer (ninthy revision of the International Classification of Diseases code 154). The median age at diagnosis of colorectal cancer in the probands was 69 years (mean, 67.2 years; SD, 12.1). The families of the 1,042 probands contained 8,738 first-degree relatives. Only 39 (0.6%) of these had been lost to follow-up. The median family size excluding the proband was 8 individuals (range, 0-24), with a mean of 8.4 individuals.

One hundred and sixty-five first-degree relatives of the 1,042 probands had been diagnosed with colorectal cancer, 156 being diagnosed at any age up to 84 years. The number of probands with zero, one, two, three, four, and five first-degree relatives with colorectal cancer was 907, 115, 14, 3, 2, and 1, respectively (corresponding numbers were 914, 110, 12, 3, 2, and 1 if restricted to age ≤84 years). Of the first-degree relatives, 66 (3.2%) parents, 89 (2.1%) siblings, and 10 (0.4%) offspring were affected with colorectal cancer (60, 86, and 10 if restricted to age ≤84 years). The median age at diagnosis of colorectal cancer was 65 years in parents (mean, 64.5 years; SD, 14.6), 62 years in siblings (mean, 60.7 years; SD, 14.9), and 39 years in offspring (mean, 42.9 years; SD, 9.5).

A total of 128 of the colorectal cancers from the 1,042 probands displayed MMR deficiency detected by MSI [12.2%; 95% confidence interval (95% CI), 10.3-14.4]. The presence of MSI was significantly correlated with an early age at diagnosis of colorectal cancer in the proband (P < 0.001). The frequency of MSI in tumors from probands ages <35, 35 to 44, 45 to 54, 55 to 64, and >65 years at diagnosis was 41.7% (5 of 12), 35.1% (13 of 37), 12.0% (13 of 108), 8.4% (18 of 213), and 11.8% (79 of 672), respectively.

Thirty-seven probands were carriers of a Mendelian predisposition to colorectal cancer (3.6%; 95% CI, 2.6-5.0): 29 (2.8%; 95% CI, 1.9-4.0) with a germ-line mutation in either MLH1 or MSH2; 2 probands had FAP (0.2%; 95% CI, 0.02-0.7); 1 had juvenile polyposis (0.1%; 95% CI, 0.002-0.5); and 5 were carriers of biallelic MYH1 variants (0.5%; 95% CI, 0.1-1.1). A further five probands were heterozygous for MYH1 variants (0.5%; 95% CI, 0.1-1.1). In addition, one proband with no family history of colorectal cancer had been documented to harbor a germ-line MSH6 mutation. The probability of identifying a germ-line mutation in either MLH1 or MSH2 was highly correlated with an early age at diagnosis of colorectal cancer: 16.7%, 29.7%, 6.5%, 3.3%, and 0.3% for probands ages <35, 35 to 44, 45 to 54, 55 to 64, and >65 years, respectively. The spectrum of mutations was restricted, one in MSH2 and two thirds of MLH1 mutations being ascribable to two founder mutations (8, 24).

**Familial colorectal cancer risks.** Risks in relatives of probands with MSI-positive cancers were significantly greater than in relatives of probands with MSI-negative cancers (log-rank test of difference: χ² = 82.49; 1 degree of freedom (df); P < 0.0001). Cox analysis with the robust variance correction confirmed the significance of MSI status on familial colorectal cancer risks (P < 0.0001). The hazard ratio for colorectal cancer associated with MSI was 3.71 (95% CI, 2.50-5.49). There was no difference in risks among parents, siblings, or offspring of probands with MSI or MS cancers (log-rank test of difference: χ² = 1.41; df, 2; P = 0.49 and χ² = 3.22; df, 2; P = 0.20, respectively). Risks at ages 40, 60, and 80 years associated with MSI and MS cancers were 1.4%, 4.5%, and 11.8% and 0.1%, 0.8%, and 3.5%, respectively (Fig. 1A). Corresponding risks of colorectal cancer in the general population at ages 40, 60, and 80 years are 0.06%, 0.5%, and 3.1%.

Figure 1B and C shows the cumulative probability of colorectal cancer in first-degree relatives of probands ages <55 and ≥55 years at diagnosis, stratified by MSI status. In both, colorectal cancer risks were significantly greater for relatives of probands with MSI cancers (log-rank test of difference: χ² = 41.97; df, 1; P < 0.0001 and χ² = 35.37; df, 1; P < 0.0001, respectively). For relatives of probands ages <55 years at diagnosis, risks in relatives at ages 40, 60, and 80 years were 4.0%, 13.0%, and 27.9% if the cancer of the proband was MSI positive and 0.3%, 2.2%, 6.1%, respectively, if the cancer of the probands was MS. In contrast, for relatives of probands ages ≥55 at diagnosis, the risks of colorectal cancer at ages 40, 60, and 80 years were 0.8%, 2.9%, and 9.2% for MSI positive and 0.1%, 0.7%, and 3.2% for MS cancers, respectively.

The familial risk of colorectal cancer in relatives of probands with MSI cancers was significantly greater in relatives of MMR gene carriers than noncarriers (log-rank test of difference: χ² = 127.2; df, 1; P < 0.0001; hazard ratio, 13.17; 95% CI, 7.23-23.98; Fig. 1D). Cumulative risks in relatives of carriers were 6.6%, 20.7%, and 37.6% at ages 40, 60, and 80 years, respectively. Corresponding risks at ages 40, 60, and 80 years in relatives of noncarriers were 0.1%, 1.0%, and 6.6%, respectively (log-rank test of difference: χ² = 3.20; df, 1; P = 0.07), comparable with that observed in relatives of probands with MS tumors.

Overall, 156 relatives were diagnosed with colorectal cancer at any age up to 84 years compared with an expected number of 90.25 based on national incidence rates (Table 1). Risks of colorectal cancer in relatives were strongly related to an early age at diagnosis of colorectal cancer in the proband. Thirty-one relatives of MSI probands were affected compared with 10.17 expected number of 83.35 (SMR, 1.50).

One hundred and five first-degree relatives of probands with MS cancers were diagnosed with colorectal cancer compared with an expected number of 6.90 (SMR, 4.50); by contrast, 125 first-degree relatives of probands were diagnosed with colorectal cancer after age 55 years with an expected number of 83.35 (SMR, 1.50).

One hundred and five first-degree relatives of probands with MS cancers were diagnosed with colorectal cancer compared with 80.07 expected (SMR, 1.31), whereas 51 first-degree relatives of MSI probands were affected compared with 10.17 expected (SMR, 5.01). These data suggest that ~62% of the excess risk (observed minus expected number) of colorectal cancer observed in first-degree relatives of probands in this series was ascribable to MSI status.

The excess risk of colorectal cancer associated with MSI status was primarily a consequence of the effect of MSH2 or MLH1 carrier status on familial risk. Thirty-four relatives of the 29 MMR mutation carriers developed colorectal cancer compared with an expected number of 1.14 relatives (SMR, 29.85), accounting for 80% of the excess risk associated with MMR.
deficiency in probands. In contrast, among relatives of probands with MSI tumors who did not possess a germ-line MMR mutation, there were 17 cases of colorectal cancer diagnosed compared with 7.97 expected cases (SRR, 2.13).

Among affected relatives of MS probands, two were first-degree relatives of probands with FAP (0.14 expected; SMR, 14.1; 95% CI, 1.71-50.1) and one was a relative of probands harboring MYH variants (0.71 expected; SMR, 1.4; 95% CI, 0.04-7.85). Taking into account these affected relatives, ~32% of the excess risk of colorectal cancer associated with family history remains unexplained by the known genes.

Complex segregation analysis was applied to the pedigrees of MS probands to investigate the genetic basis of the residual familial aggregation of colorectal cancer. The families of probands with FAP and MYH variants were excluded from this analysis to further minimize the effect of these mutations on the familial colorectal cancer. A model not providing for family resemblance, in which the familial occurrence of colorectal cancer can be attributed to chance (i.e., sporadic), could be rigorously rejected ($P < 10^{-3}$). The various models incorporating a single major gene locus ($q>0$) were, however, indistinguishable. A model based on polygenic heritability provided for a marginally better explanation of the residual familial aggregation of colorectal cancer than a major gene model (Akaike’s information criterion, 1,086 compared with 1,090).

**Discussion**

A major advantage of the present design to estimate familial risks was the use of information derived from the centralized health care system, as well as the national population and cancer registers, to confirm cancer diagnoses in the cohort. In this way, potential information bias, particularly measurement error and recall bias inherent in relying on probands as the only source of family history information, has been limited.

Choice of an appropriate comparison group is an important issue in the design and interpretation of epidemiologic studies of this nature. In our study, age-, sex-, and calendar-specific incidence rates for Finland were used for comparison to minimize bias. Survivorship is a potential source of bias in the ascertainment of family histories through probands, although recent data suggest that this is unlikely to be a real concern at least with respect to germ-line MMR defects (37). Although we sampled a high proportion of all colorectal cancers, approximately two thirds of the target population...
Although we did not systematically assay for mutations in other MMR genes, it is noteworthy that one individual without a family history of colorectal cancer was documented to be a carrier of a mutation in MSH6.

The overall contribution of MSH2 and MLH1 mutations to colorectal cancer incidence in this series was 2.8%. Other studies have reported lower estimates for the prevalence of MMR mutations in unselected colorectal cancer cases (38). There are a restricted number of founder mutations in Finland that may account for the higher frequency of mutations detected in this ethnic group than that reported in analyses of some highly out-bred populations. It is therefore conceivable that in the mutations of the different population in MMR genes other than MSH2 and MLH1 may have a higher contribution to the overall excess risk associated with MSI status.

An important conclusion from our study is that a third of the excess risk of colorectal cancer associated with family history is not explained by mutations in known genes. Although this estimate does not take into account the possible contribution of attenuated colorectal cancer predisposition syndromes (such as attenuated FAP) to familial risk, the contribution of such disorders is unlikely to significantly affect the estimate. Similarly, restricting the mutation analysis of MYH to the common Y165C and G382D variants in the primary screen is unlikely to have affected our findings as heterozygote carriers were screened for rare disease-causing variants.

The existence of Amsterdam-positive HNPCC families with no DNA MMR defect recently termed “familial colorectal cancer type X” (39) provides evidence for a dominant model as the mechanism of residual susceptibility. In contrast, a recent segregation analysis of colorectal cancer families, in which the mechanism of residual susceptibility. In contrast, a recent segregation analysis of colorectal cancer families, in which MSH2/MLH1 mutations were detected in the primary screen is unlikely to significantly affect the estimate. Similarly, restricting the mutation analysis of MYH to the common Y165C and G382D variants in the primary screen is unlikely to have affected our findings as heterozygote carriers were screened for rare disease-causing variants.

The existence of Amsterdam-positive HNPCC families with no DNA MMR defect recently termed “familial colorectal cancer type X” (39) provides evidence for a dominant model as the mechanism of residual susceptibility. In contrast, a recent segregation analysis of colorectal cancer families, in which MSH2/MLH1 mutations had been excluded, proposed recessive susceptibility as a basis of the residual genetic risk (40). The analysis did not, however, incorporate MSI status.

| Table 1. Colorectal cancer incidence up to age 85 years in first-degree relatives of colorectal cancer probands (expected figures calculated from national rates from Finland) |
|-----------------|--------|-----------------|--------|-----------------|--------|
| Age of relative (y) | O | E | SMR (95% CI) | O | E | SMR (95% CI) | O | E | SMR (95% CI) |
| Age of proband (y) | <55 | 55+ | All | <55 | 55+ | All | <55 | 55+ | All | <55 | 55+ | All |
| All | 17 | 1.62 | 10.47 | (6.10-16.77) | 33 | 14.99 | 2.20* | (1.52-3.09) | 50 | 16.61 | 3.01* | (2.23-3.97) |
| MS | 4 | 1.32 | 3.04 | (0.83-7.79) | 16 | 13.29 | 1.20 | (0.69-1.96) | 20 | 14.61 | 1.37 | (0.84-2.11) |
| MSH2 | 13 | 5.56 | 2.34 | (1.24-4.00) | 92 | 74.51 | 1.23 | (1.00-1.51) | 105 | 80.07 | 1.31 | (1.07-1.59) |
| MLH1 | 13 | 0.31 | 42.20 | (22.48-72.17) | 17 | 1.70 | 9.98* | (5.82-15.98) | 30 | 2.01 | 14.98* | (9.58-21.30) |
| MMR carriers | 18 | 1.34 | 13.45 | (7.97-21.26) | 33 | 8.84 | 3.74* | (2.57-5.25) | 51 | 10.17 | 5.01* | (3.73-6.59) |

NOTE: SMRs with 95% CI in first-degree relatives of all probands, probands with MMR-deficient cancer, and probands possessing germ-line MSH2 or MLH1 mutations. Abbreviations: O, observed; E, expected.

* P < 0.001.
† P < 0.01.
or take into account MYH. Although the nature of the residual inherited susceptibility to colorectal cancer is at present not fully defined, a model in which high-risk alleles account for all of the excess familial risk therefore seems unlikely. Indeed, in our analysis, a simple polygenic heritability provided a marginally better explanation of the residual clustering of colorectal cancer after excluding involvement of the known loci.

Stratifying cancers based on MSI status provides a powerful method of delineating familial cases with different risk profiles; moreover, germ-line MMR gene mutations make the largest contribution to the familial risk, particularly at young ages. The observation that a third of the excess familial risk is unaccounted for by known genes provides a strong rationale for directing research to families segregating MS cancers to identify novel susceptibility genes. Findings also indicate that screening individuals with a family history of colorectal cancer should not be restricted to those at risk of the known Mendelian forms of susceptibility, but screening programs should take into account the different risk profiles.

**References**


**Acknowledgments**

We thank Sini Martenen for her diligent pedigree work.
Explaining the Familial Colorectal Cancer Risk Associated with Mismatch Repair (MMR)-Deficient and MMR-Stable Tumors


Updated version  Access the most recent version of this article at:  
[http://clincancerres.aacrjournals.org/content/13/1/356](http://clincancerres.aacrjournals.org/content/13/1/356)

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.