Research Article

Low-Penetrance Susceptibility Variants in Familial Colorectal Cancer

Iina Niittymäki1, Eevi Kaasinen1, Sari Tuupanen1, Auli Karhu1, Heikki Järvinen2, Jukka-Pekka Mecklin3, Ian P.M. Tomlinson4, Maria Chiara Di Bernardo5, Richard S. Houlston5, and Lauri A. Aaltonen1

Abstract

Background: Genomewide association studies have identified 10 low-penetrance loci that confer modestly increased risk for colorectal cancer (CRC). Although they underlie a significant proportion of CRC in the general population, their impact on the familial risk for CRC has yet to be formally enumerated. The aim of this study was to examine the combined contribution of the 10 variants, rs6983267, rs4779584, rs4939827, rs16892766, rs10795668, rs3802842, rs4444235, rs9929218, rs10411210, and rs961253, on familial CRC.

Methods: The population-based series of CRC samples included in this study consisted of 97 familial cases and 691 sporadic cases. Genotypes in the 10 loci and clinical data, including family history of cancer verified from the Finnish Cancer Registry, were available. The overall number of risk alleles (0-20) was determined, and its association with familial CRC was analyzed. Excess familial risk was estimated using cancer incidence data from the first-degree relatives of the cases.

Results: A linear association between the number of risk alleles and familial CRC was observed ($P = 0.006$). With each risk-allele addition, the odds of having an affected first-degree relative increased by 1.16 (95% confidence interval, 1.04-1.30). The 10 low-penetrance loci collectively explain $\sim 9\%$ of the variance in familial risk for CRC.

Conclusions: This study provides evidence to support the previous indirect estimations that these low-penetrance variants account for a relatively small proportion of the familial aggregation of CRC.

Impact: Our results emphasize the need to characterize the remaining molecular basis of familial CRC, which should eventually yield in individualized targeting of preventive interventions.

Cancer Epidemiol Biomarkers Prev; 19(6); 1478–83. ©2010 AACR.

Introduction

Family history is a major risk factor for colorectal cancer (CRC), with individuals with an affected first-degree relative having a 2-fold increased risk (1). Concordance between monozygotic and dizygotic twins indicate hereditary factors to underlie $\sim 35\%$ of CRC (2). However, germline mutations in high-penetrance genes, including MLH1, MSH2, APC, and MYH account for only $\sim 5\%$ of all CRCs (3). According to the "common disease–common variant" concept, much of the remaining inheritance is likely caused by a large number of low-penetrance variants that are common in the population. Genomewide association studies conducted in United Kingdom and Canada, in which large case-control sets have been genotyped for single-nucleotide polymorphisms, have thus far identified 10 chromosomal loci that confer a modest risk for CRC. The first CRC-associating variant to emerge in 2007 was rs6983267 at 8q24, which increases the risk for CRC with an odds ratio of 1.21 [95% confidence interval (95% CI), 1.15-1.27; refs. 4-6]. Subsequently, five more loci at 18q21, 15q13, 8q23, 10p14, and 11q23 were published, in which the most strongly associating single-nucleotide polymorphisms were rs4939827, rs4779584, rs16892766, rs10795668, and rs3802842, respectively (7-11). In the first phases of the genomewide association studies, <10,000 cases and controls were genotyped with microarrays containing half a million single-nucleotide polymorphisms, but all the six associations were later replicated in around 7,000 to 15,000 cases and controls. After the first round, additional four variants, rs4444235 at 14q22, rs9929218 at 16q22, rs10411210 at 19q13, and rs961253 at 20p12, were identified through meta-analysis in which

Authors' Affiliations: 1Department of Medical Genetics, Genome-Scale Biology Research Program, Biomedicum Helsinki, University of Helsinki; 2Department of Surgery, Helsinki University Hospital, Helsinki, Finland; 3Department of Surgery, Jyväskylä Central Hospital, Jyväskylä, Finland; 4Molecular and Population Genetics, Wellcome Trust Centre for Human Genetics, Oxford, United Kingdom; and 5Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom

Corresponding Author: Lauri A. Aaltonen, Department of Medical Genetics, Biomedicum Helsinki, University of Helsinki, P.O. Box 63, FIN-00014 Helsinki, Finland. Phone: +358-9-19125595; Fax: +358-9-19125105. E-mail: lauri.aaltonen@helsinki.fi

doi: 10.1158/1055-9965.EPI-09-1320

©2010 American Association for Cancer Research.
these associations were replicated in more than 20,000 cases and controls (12).

All the 10 low-penetration loci independently predispose to CRC with allelic odds ratios of 1.10 to 1.26 (95% CI, 1.06-1.34) and have risk allele frequencies of 0.07 to 0.90 in the general population (13). Although, typically, the associating tagging single-nucleotide polymorphisms are merely correlated with the actual causal variants, we have shown that rs6983267 at 8q24 affects Wnt signaling by disrupting an enhancer element and possibly driving the expression of the 335 kb distal MYC oncogene (14). In 18q21, the functional change is a novel variant altering SMAD7 expression (15). Indeed, all the 10 tagging single-nucleotide polymorphisms locate in either intergenic or intronic areas of the genome and many might therefore affect gene expression through distant regulatory elements. Aberrant transforming growth factor β superfamily signaling seems to be at the biological basis of several loci because genes such as SMAD7, GREM1, BMP2, BMP4, and RHPN2 are inside or near the CRC-associating linkage disequilibrium regions (13).

Although each of the 10 variants independently confer only modest predisposition to CRC, their additive contribution on an individual’s risk can be much higher. It has been estimated that individuals with ≥15 risk alleles in the 10 loci are at 3.6-fold risk for developing CRC compared with individuals possessing nine risk alleles (12). Clearly, the individuals with many risk alleles can be predicted to display positive family history for CRC more often than others. However, ideal data sets, population-based material with full cancer information on family members, to examine the size of the effect are few. Recently, Middeldorp et al. (16) studied six of the loci (at 8q24, 15q13, 8q23, 10p14, and 11q23) in a Dutch CRC cohort, concluding that the risk alleles were indeed enriched in the familial cases. The Dutch cohort consisted of 995 individuals that were selected based on early onset of the disease or for having at least two affected first-degree relatives. Middeldorp et al. (16) showed that the patients with a family history of CRC harbored significantly more risk alleles in the six loci than the early-onset solitary cases.

In this study, we have sought to establish the combined impact of the 10 low-penetration variants on the familial aggregation of CRC in an unbiased population-based series of 826 Finnish CRC cases, with verified cancer data from all first-degree relatives.

**Materials and Methods**

**Patient samples**

A previously characterized population-based sample series of 1042 CRCs collected between 1994 and 1998 from nine central hospitals in south-eastern Finland was used in this study (17, 18). As described in the previous studies, DNA was extracted from fresh-frozen normal tissue or blood with standard methods. Comprehensive registry-based clinical data and tumor characteristics evaluated by a pathologist were collected from the patients. Information on cancer occurrence in the first-degree relatives (parents, siblings, and offspring) of the 1,042 probands was obtained from the cancer registry, death certificates, and medical records. Microsatellite instability status was determined for all the tumors, and subsequently, the patients with microsatellite instability tumors were screened for mutations in MLH1 and MSH2 genes to identify Lynch syndrome patients. Diagnoses of other high-penetration syndromes, including MYH-associated polyposis, familial adenomatous polyposis, and juvenile polyposis, were also established based on clinical features and mutational analysis (3). Samples and clinicopathologic information were obtained with informed consent and ethical review board approval in accordance with the tenets of the declaration of Helsinki.

**Genotyping**

All of the 10 single-nucleotide polymorphisms included in this study had previously been genotyped in the given sample series. Briefly, rs6983267, rs4779584, and rs4939827 were genotyped by direct genomic sequencing using Applied Biosystems BigDye v3.1 sequencing chemistry and ABI3730 Automatic DNA sequencer (12, 19). Competitive allele-specific PCR KASPar chemistry (KBiosciences Ltd.) was used for genotyping of rs3802842, rs16892766, rs10795668, rs4444235, rs9929218, rs10411210, and rs961253 (9, 11).

Apart from the 773 cases of which all of the 10 single-nucleotide polymorphisms had successfully been genotyped, there were 60 cases in which one of the genotypes was missing because of a failure in the analysis. The rerun was feasible for 53 samples that were sequenced for the missing single-nucleotide polymorphism locus, resulting in 826 successfully genotyped individuals. The additional sequencing created no bias because all of the samples with one missing genotype were included in the rerun, irrespective of how many risk alleles the individuals were known to harbor. Genomic DNA extracted from either fresh-frozen normal tissue or blood was amplified by standard PCR using AmpliTaq Gold DNA polymerase (Applied Biosystems). Primers used in the PCRs were designed with Primer3 software (20), based on reference sequences from the Ensembl database (21). PCRs were purified with ExoSAP-IT enzyme (USB Corporation), and sequencing was done with Applied Biosystems BigDye v3.1 sequencing chemistry and ABI3730 Automatic DNA sequencer.

**Statistical analysis**

Odds ratios, 95% CI, and P values for the associations of the 10 single-nucleotide polymorphisms with CRC were calculated with Pearson χ² test by comparing the allele frequencies between cases and controls. Information was combined from multiple single-nucleotide polymorphisms using an allele count model summing
the number of risk alleles (0-20) carried by each individual. This assumes that each of the alleles has an equal and additive effect on CRC risk. The associations of these numbers with familial CRC were then analyzed using logistic regression. In the regression model, the relationship of the risk allele numbers as a continuous predictor variable was analyzed with familial or sporadic CRC as a binary outcome variable. Allele groups below five were excluded from the analysis because of the small number of cases (two individuals with three risk alleles and four with four risk alleles). The effect of potential confounders, age at diagnosis, and gender was examined, analyzing them as covariables with risk alleles. Adjustment for age at diagnosis improved model fit and was therefore included in analyses. This was categorized in 10-year intervals (0-39, 40-49, 50-59, 60-69, 70-79, 80-) and considered as a factor variable in the model. The model fit was evaluated by comparing null deviance with residual deviance. The commonest risk allele number in the entire sample series was 10, which was used as a reference group (odds ratio, 1.00) when calculating odds ratios. The association of risk allele numbers with other familial cancers and several tumor characteristics (microsatellite instability status, Duke’s stage, histologic grade, and location) was also tested. Pearson χ² test with Yates’s continuity correction was used to calculate potential differences in clinical and demographic characteristics between the cases with familial and sporadic CRC. R software (version 2.10.0) was used in all the above analyses.

Table 1. Summary of the low-penetrance CRC variants in the Finnish population

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Chr location</th>
<th>Risk allele</th>
<th>Risk allele frequency*</th>
<th>Cases</th>
<th>Controls</th>
<th>Allelic OR† (95% CI)</th>
<th>P</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6983267</td>
<td>8q24</td>
<td>G</td>
<td>0.53</td>
<td>996</td>
<td>1,012</td>
<td>1.22 (1.08-1.38)</td>
<td>0.002</td>
<td>Tuupanen et al. (19)</td>
</tr>
<tr>
<td>rs4779584</td>
<td>15q13</td>
<td>T</td>
<td>0.3</td>
<td>981</td>
<td>1,024</td>
<td>1.14 (1.00-1.31)</td>
<td>0.053</td>
<td>Houlston et al. (12)</td>
</tr>
<tr>
<td>rs4939827</td>
<td>18q21</td>
<td>T</td>
<td>0.46</td>
<td>970</td>
<td>969</td>
<td>1.20 (1.06-1.36)</td>
<td>0.004</td>
<td>Houlston et al. (12)</td>
</tr>
<tr>
<td>rs16892766</td>
<td>8q23</td>
<td>C</td>
<td>0.12</td>
<td>968</td>
<td>1,003</td>
<td>1.21 (1.01-1.46)</td>
<td>0.041</td>
<td>Tomlinson et al. (9)</td>
</tr>
<tr>
<td>rs10795668</td>
<td>10p14</td>
<td>G</td>
<td>0.69</td>
<td>951</td>
<td>983</td>
<td>1.19 (1.04-1.37)</td>
<td>0.014</td>
<td>Tomlinson et al. (9)</td>
</tr>
<tr>
<td>rs3802842</td>
<td>11q23</td>
<td>C</td>
<td>0.24</td>
<td>970</td>
<td>994</td>
<td>1.21 (1.05-1.40)</td>
<td>0.008</td>
<td>Pittman et al. (11)</td>
</tr>
<tr>
<td>rs4444235</td>
<td>14q22</td>
<td>C</td>
<td>0.43</td>
<td>933</td>
<td>828</td>
<td>1.16 (1.02-1.33)</td>
<td>0.028</td>
<td>Houlston et al. (12)</td>
</tr>
<tr>
<td>rs9929218</td>
<td>16q22</td>
<td>G</td>
<td>0.77</td>
<td>945</td>
<td>836</td>
<td>1.03 (0.88-1.20)</td>
<td>0.751</td>
<td>Houlston et al. (12)</td>
</tr>
<tr>
<td>rs10411210</td>
<td>19q13</td>
<td>C</td>
<td>0.84</td>
<td>917</td>
<td>795</td>
<td>1.18 (0.97-1.42)</td>
<td>0.091</td>
<td>Houlston et al. (12)</td>
</tr>
<tr>
<td>rs961253</td>
<td>20p12</td>
<td>A</td>
<td>0.29</td>
<td>941</td>
<td>814</td>
<td>1.23 (1.07-1.42)</td>
<td>0.005</td>
<td>Houlston et al. (12)</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single-nucleotide polymorphism; Chr, chromosome; OR, odds ratio.
*In our controls.
†Risk allele versus neutral allele.

Table 2. Characteristics of 788 analyzed CRC cases

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Familial (n = 97)</th>
<th>Sporadic (n = 691)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥69</td>
<td>59</td>
<td>351</td>
<td>0.08</td>
</tr>
<tr>
<td>&lt;69</td>
<td>38</td>
<td>340</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>52</td>
<td>354</td>
<td>0.74</td>
</tr>
<tr>
<td>F</td>
<td>45</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>MSI status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSI</td>
<td>11</td>
<td>76</td>
<td>0.94</td>
</tr>
<tr>
<td>MSS</td>
<td>86</td>
<td>615</td>
<td></td>
</tr>
<tr>
<td>Duke’s stage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-B</td>
<td>62</td>
<td>398</td>
<td>0.25</td>
</tr>
<tr>
<td>C-D</td>
<td>34</td>
<td>291</td>
<td></td>
</tr>
<tr>
<td>Histologic grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>84</td>
<td>625</td>
<td>0.64</td>
</tr>
<tr>
<td>3-4</td>
<td>9</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Location*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distal</td>
<td>62</td>
<td>476</td>
<td>0.5</td>
</tr>
<tr>
<td>Proximal</td>
<td>33</td>
<td>211</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Some of the numbers do not match do to missing data.
Abbreviations: M, male; F, female; MSI, microsatellite instability; MSS, microsatellite-stable.
*Distal, from splenic flexure to rectum; proximal, from cecum to transverse colon.
with the 10 risk loci by Poisson regression. These analyses were conducted using the statistical software program STATA (version 10; Stata Corporation).

Results

The tagging single-nucleotide polymorphisms analyzed were rs6983267 at 8q24, rs4779584 at 15q13, rs4939827 at 18q21, rs16892766 at 10q14, rs3802842 at 11q23, rs4444235 at 14q22, rs9929218 at 16q22, rs10411210 at 19q13, and rs961253 at 20p12. Our sample series of ~1,000 CRC cases and population-matched controls was included in the replication phase of the United Kingdom genomewide association study and in the meta-analysis (9, 11, 12). Results from the genotyping studies on these Finnish samples are summarized in Table 1.

Altogether, 826 individuals from a population-based CRC series had been or were successfully genotyped for all the 10 variants. Of the 826 cases, 37 were known to carry germline mutations (in either MLH1, MSH2, MSH6, MYH, APC, or ALK3) and were therefore excluded from the analysis. Of the remaining 789 patients, 97 had a family history of CRC (at least one affected first-degree relative) and 691 were sporadic cases. One individual was excluded from the analysis because of non-Finnish descent and missing family history records. There were altogether 426 probands with other cancers than CRC in the first-degree relatives. Detailed demographic and clinicopathologic characteristics of the 97 familial and 691 sporadic cases included in this study are reported in Table 2. Age at diagnosis in the familial cases was slightly higher than in the sporadic cases (69.8 and 67.8 y, respectively; \( P = 0.08 \)). This could not be attributed to any effects of the low-penetrance variants because the median number of risk alleles was 10 in both younger (<69 y) and older cases (≥69 y). There were no significant difference between the familial and sporadic cases with respect to gender (\( P = 0.74 \)), microsatellite instability status (\( P = 0.94 \)), Duke’s stage (\( P = 0.25 \)), histologic grade (\( P = 0.64 \)), or tumor location (\( P = 0.50 \); Table 2).

The overall number of risk alleles (1-20) in the 10 single-nucleotide polymorphisms was calculated for each of the 788 cases. Figure 1 shows the distribution of the number of risk alleles in familial and sporadic cases. The median number of risk alleles was 10 in both younger (<69 y) and older cases (≥69 y). There were no significant difference between the familial and sporadic cases with respect to age at diagnosis (\( P = 0.08 \)). This could not be attributed to any effects of the low-penetrance variants because the median number of risk alleles was 10 in both younger (<69 y) and older cases (≥69 y). There were no significant difference between the familial and sporadic cases with respect to gender (\( P = 0.74 \)), microsatellite instability status (\( P = 0.94 \)), Duke’s stage (\( P = 0.25 \)), histologic grade (\( P = 0.64 \)), or tumor location (\( P = 0.50 \); Table 2).

The overall number of risk alleles (1-20) in the 10 single-nucleotide polymorphisms was calculated for each of the 788 cases. Figure 1 shows the distribution of the number of risk alleles in familial and sporadic cases. The median number of risk alleles in familial cases was 11 compared with 10 for sporadic cases. There was a linear association between the number of risk alleles and familial CRC (\( P_{Trend} = 0.006 \)). For each one risk allele increase, the odds of having familial cancer increased by a factor of 1.16 (95% CI, 1.04-1.30; \( P = 0.006 \)).

<table>
<thead>
<tr>
<th>No. of risk alleles</th>
<th>Familial cases, n (%)</th>
<th>Sporadic cases, n (%)</th>
<th>OddsA</th>
<th>OR if 10 (ref)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3 (3.1)</td>
<td>4 (0.6)</td>
<td>0.2</td>
<td>2.15</td>
</tr>
<tr>
<td>14</td>
<td>4 (4.1)</td>
<td>15 (2.2)</td>
<td>0.18</td>
<td>1.84</td>
</tr>
<tr>
<td>13</td>
<td>1 (1.0)</td>
<td>44 (6.4)</td>
<td>0.15</td>
<td>1.58</td>
</tr>
<tr>
<td>12</td>
<td>16 (16.5)</td>
<td>69 (10.0)</td>
<td>0.13</td>
<td>1.36</td>
</tr>
<tr>
<td>11</td>
<td>27 (27.8)</td>
<td>121 (17.5)</td>
<td>0.11</td>
<td>1.17</td>
</tr>
<tr>
<td>10</td>
<td>12 (12.4)</td>
<td>115 (16.6)</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>13 (13.4)</td>
<td>135 (19.5)</td>
<td>0.08</td>
<td>0.86</td>
</tr>
<tr>
<td>8</td>
<td>14 (14.4)</td>
<td>100 (14.5)</td>
<td>0.07</td>
<td>0.74</td>
</tr>
<tr>
<td>7</td>
<td>5 (5.2)</td>
<td>49 (7.1)</td>
<td>0.06</td>
<td>0.63</td>
</tr>
<tr>
<td>6</td>
<td>1 (1.0)</td>
<td>23 (3.3)</td>
<td>0.05</td>
<td>0.54</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>11 (1.6)</td>
<td>0.04</td>
<td>0.47</td>
</tr>
<tr>
<td>Total</td>
<td>97</td>
<td>691</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: A, adjusted with age at diagnosis; ref, reference.

*Odds ratio if 10 risk alleles (the most common number in all cases) set as a reference group.
10 risk alleles, which was the most common number of risk alleles in the entire sample series. The individuals with 15 risk alleles were more than twice as likely to have affected first-degree relatives compared with the individuals possessing 10 risk alleles (Table 3).

The association of risk allele numbers with other familial cancers and several tumor characteristics, including microsatellite instability status, Duke’s class, histologic grade, and location, was examined. No relationship between the number of risk alleles and other familial cancers than CRC was found (odds ratio, 1.06; 95% CI, 0.99–1.14; \( P = 0.10 \)). A trend toward the association of increasing numbers of risk alleles with distal or left side, rather than proximal or right side, location of the tumor was observed (odds ratio, 1.08; 95% CI, 1.00–1.17; \( P = 0.04 \)) but not when based on comparison of colonic versus rectal disease (\( P = 0.23 \)). There was no association between the risk allele numbers and any other tumor characteristics: microsatellite instability status (odds ratio, 0.93; 95% CI, 0.83–1.04; \( P = 0.23 \)), Duke’s class (odds ratio, 1.03, 95% CI, 0.96–1.11; \( P = 0.40 \)), or histologic grade (odds ratio, 0.91; 95% CI, 0.80–1.04; \( P = 0.18 \)).

Finally, the excess familial risk explained by the 10 low-penetration variants was calculated. This estimates which proportion of the familial aggregation of CRC in our sample series can be attributed to the given loci. Altogether, 84 first-degree relatives were diagnosed with CRC by the age of 80 years compared with 60.3 expected. On the basis of Poisson regression analysis of the observed CRC in first-degree relatives accounting for expected numbers, the 10 loci were calculated to be responsible for 8.7% of the total variance in familial CRC risk (upper 95% CI, 19.0%).

**Discussion**

Excluding the high-penetration CRC syndromes, ~10% of the CRC patients have a family history of the disease (3). This nonsyndromic familial CRC can be defined as an apparently sporadic form of the disease that occurs in families more often than expected by chance. Several common low-penetration variants are likely to contribute to this predisposition, and their identification may enable genetic risk profiling and tailoring of preventive interventions. Although empirical risk stratification based on family history of CRC is already feasible, it does not fully address differences in the inheritance of risk alleles among offspring and hence different risk profiles.

In this study, we analyzed the contribution of all the recently identified 10 low-penetrance CRC variants to the familial aggregation in a Finnish population-based sample series. Results of the genomewide association study’s replication in the Finnish sample series show that most of the 10 associations are well replicated in our sample series.

In our study, it is striking that the impact of the 10 variants on the inherited predisposition to CRC is apparent despite its modest size. We observed a clear association of increasing numbers of risk alleles in the 10 loci with familial CRC, compatible with the observation of Middeldorp et al. (16) from six of the loci (16). The 995 Dutch cases were selected based on family history or early onset of CRC, whereas our sample series was population based. Middeldorp et al. (16) also reported an association of risk allele number with early onset compared with late-onset disease, which was not observed in our study.

Three low-penetration loci have been reported to be more common in rectal than in colonic tumors: rs3802842 at 11q23, rs4939827 at 18q21, and rs1079568 at 10p14 (9, 10). We also saw a trend toward the association of increasing numbers of risk alleles and odds of having distal rather than proximal cancer (\( P = 0.04 \)). No enrichment of risk alleles was detected in any subgroup in terms of age at diagnosis, gender, Duke’s stage, histologic grade, or microsatellite instability status of the tumor. This suggests that, in combination, the 10 low-penetration variants do not clearly contribute to a specific type of CRC but have a generic influence. We did not find any association between increasing numbers of risk alleles in the 10 loci and odds of having other familial cancer. This was expected because pleiotropic effects are reported only in the 8q24 CRC locus that increases the risk for prostate and ovarian cancers (6, 23).

In a meta-analysis of the UK genomewide association studies, the excess familial risk for CRC attributable to the 10 low-penetration variants was calculated to be 6% (12). This was, however, an indirect estimate based on a log-additive model. To address the contribution of the 10 loci to the familial risk for CRC, we directly calculated the excess familial risk using incidence data from the first-degree relatives of the genotyped probands. Although a limitation of this study is its relatively small sample size, this is the first time the excess risk associated with the 10 variants has been estimated using comprehensive registry-based family histories in a well-characterized, population-based sample series. This is a major advantage because self-reported family histories tend to be inaccurate. Furthermore, we used age, sex, and calendar period incidence rates for the general population of Finland for comparison when calculating the expected incidences in the first-degree relatives of the genotyped probands, which minimizes bias. Our observation that ~9% of the variance in familial risk is explained by the 10 loci provides evidence to support the indirect estimates that most the inherited predisposition to CRC remains unaccounted for by the currently known variants.

As new low-penetration variants are identified in even larger genomewide association studies and meta-analyses, their role in explaining the familial risk for CRC is likely to be better refined. Tenesa and Dunlop (2009) have estimated that up to ~170 common variants could independently contribute to the observed hereditary predisposition to CRC (13). On the other hand, the existing genomewide association studies have captured most of the common variation in European populations, and hence, finding high numbers of additional low-penetrance
Low-Penetrance Variants and Familial Colorectal Cancer

variants seems unlikely. To facilitate the characterization of CRC predisposition, an international consortium, COGENT (COlorectal cancer GENeTics), has been established, which currently has access to >48,000 cases and 43,000 controls from 20 different research groups all over the world (24).

An important part of the familial risk might also be explained by moderate-penetrance loci because these have not been well captured by the current genotyping platforms. Next-generation sequencing efforts are expected to play a crucial role in discovering these variants. Hemminki et al. (25) have suggested that many of the low-penetrance associations could actually be markers of rarer functional alleles and hence explain a larger part of the excess familial risk than currently estimated. However, this is not supported by the causal variant in 8q24, being the tagging single-nucleotide polymorphism itself or by the 18q21 locus, in which the causal variant has a frequency of 43% and confers a similar increase in risk (odds ratio, 1.22) than the tagging single-nucleotide polymorphism (14, 15). It remains to be established whether the functional changes in any of the remaining eight loci are moderate-penetrance variants.

In summary, by statistically analyzing the currently known 10 low-penetrance CRC variants in Finland, we have shown a significant association between increasing numbers of risk alleles in the 10 loci and odds of having familial rather than sporadic CRC. Using registry-based family history data, we estimated the 10 loci to underlie ~9% of the variance in familial risk. This study contributes to the understanding of the genetic landscape of CRC but also emphasizes the need to characterize the remaining inherited predisposition to this malignancy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Sini Marttinen for managing the patient data.

Grant Support

Academy of Finland (Centre of Excellence in Translational Genome-Scale Biology grant 6302352), the Finnish Cancer Society, the Sigrid Juselius Foundation, the Paulo Foundation (I. Niittymäki), and Cancer Research UK (C1298/A8362 supported by the Bobby Moore Fund; R.S. Houlston). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 12/31/2009; revised 03/18/2010; accepted 04/01/2010; published OnlineFirst 05/25/2010.

References


23. Ghousaini M, Song H, Koessler T, et al. Multiple loci with different functional changes in any of the remaining eight loci are moderate-penetrance variants.

In summary, by statistically analyzing the currently known 10 low-penetrance CRC variants in Finland, we have shown a significant association between increasing numbers of risk alleles in the 10 loci and odds of having familial rather than sporadic CRC. Using registry-based family history data, we estimated the 10 loci to underlie ~9% of the variance in familial risk. This study contributes to the understanding of the genetic landscape of CRC but also emphasizes the need to characterize the remaining inherited predisposition to this malignancy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Sini Marttinen for managing the patient data.

Grant Support

Academy of Finland (Centre of Excellence in Translational Genome-Scale Biology grant 6302352), the Finnish Cancer Society, the Sigrid Juselius Foundation, the Paulo Foundation (I. Niittymäki), and Cancer Research UK (C1298/A8362 supported by the Bobby Moore Fund; R.S. Houlston). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 12/31/2009; revised 03/18/2010; accepted 04/01/2010; published OnlineFirst 05/25/2010.
Low-Penetrance Susceptibility Variants in Familial Colorectal Cancer

Iina Niittymäki, Eevi Kaasinen, Sari Tuupanen, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-09-1320

Cited articles
This article cites 23 articles, 8 of which you can access for free at:
http://cebp.aacrjournals.org/content/19/6/1478.full.html#ref-list-1

Citing articles
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
/content/19/6/1478.full.html#related-urls

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.