MLH1 Polymorphisms and Cancer risk: a Meta-analysis Based on 33 Case-control Studies

Jia-Li Xu, Zhi-Qiang Yin, Ming-De Huang, Xie-Feng Wang, Wen Gao, Ling-Xiang Liu, Rong-Sheng Wang, Pu-Wen Huang, Yong-Mei Yin, Ping Liu, Yong-Qian Shu

Abstract

Objective: Cumulative evidence suggests that MLH1, the key component in the mismatch pathway, plays an important role in human cancers. Two potential functional polymorphisms (-93G>A and I219V) of MLH1 have been implicated in cancer risk. The aim of this meta-analysis was to summarize the evidence for associations.

Methods: Eligible studies were identified by searching the electronic literature PubMed, ScienceDirect and Embase databases for relevant reports and bibliographies. Studies were included if of case-control design investigating MLH1 polymorphisms (-93G>A and I219V) and cancer risk with sufficient raw data for analysis. Odds ratios (OR) and 95% confidence intervals (95% CI) were used to evaluate the strength of associations.

Results: Our meta-analysis from 33 published case-control studies showed the variant A allele of -93G>A polymorphism to be associated with increased risk in all genetic models (AA vs. GG: OR = 1.22, 95% CI: 1.03-1.44), especially among non-Asians (AA vs. GG: OR = 1.28, 95% CI: 1.04-1.58). For the I219V polymorphism, however, there was no main effect associated with overall cancer risk in any genetic model.

Conclusions: The meta-analysis suggested that the MLH1 -93G>A polymorphism may be a biomarker of cancer susceptibility. Large sample association studies and assessment of gene-to-gene as well as gene-to-environment interactions are required to confirm these findings.

Keywords: MLH1 - cancer - polymorphism - meta-analysis
region. Functional study by Perera et al. suggests that the -93G>A polymorphism modifies the efficiency of MLH1 transcription (Perera et al., 2011). Several studies have reported an association between this polymorphism and cancer risk, although the conclusions remain controversial. The nonsynonymous coding single nucleotide polymorphism I219V (rs1799977) in MLH1, which is located in exon 8 at nucleotide position 655, was shown to be associated with an increased risk of childhood acute lymphoblastic leukemia (Mathonnet et al., 2003) and breast cancer (Listgarten et al., 2004). It is also considered to be linked with another downstream SNP (such as IVS14-19A>G) and may have an impact on MLH1 gene function (Hutter et al., 2000).

In consideration of the important role of MLH1 in the carcinogenic process, we carried out a meta-analysis to estimate the overall cancer susceptibility with two polymorphisms of MLH1, -93G>A and I219V.

Materials and Methods

Identification and Eligibility of Relevant Studies

We attempted to includ all the case-control studies of cancer with sufficient genotyping data for at least one of the two polymorphisms, -93G>A and I219V. Eligible studies were identified by searching the electronic literature PubMed, ScienceDirect and Embase databases for relevant reports (last search update Dec 31, 2011, covering the terms ‘MLH1’, ‘polymorphism’, ‘cancer’). We also checked the references of the selected articles to identify additional eligible publications and contacted with the authors if necessary. No language restrictions were imposed.

Studies were included if they fulfilled the following criteria: (1) designed as a case-control study, (2) evaluated the association between MLH1 -93G>A or I219V polymorphism and malignancy, (3) provided raw data of individual genotype (heterozygous and homozygous of variant-type and wild-type). If studies had partly overlapped subjects, the one with a smaller sample size was eliminated (Samowitz et al., 2008; Campbell et al., 2009). One study was eliminated because it did not have an appropriate case-control design (Berndt et al., 2007). Five studies were excluded because they did not present detailed genotyping information for the two polymorphisms (Hutter et al., 2002; Burmester et al., 2004; Schafmayer et al., 2007; Allan et al., 2008; Song et al., 2010). Hence, the data for this analysis were available from 24 case-control studies, including 24, 137 cancer cases and 21,530 controls for -93G>A (from 20 studies), 10,632 cancer cases and 11,231 controls for I219V (from 18 studies).

Data Extraction

Data was extracted by two independent investigators using the same standardized form. Discrepancies were settled by reviewed again until a consensus was reached. The following information was sought from each study: author, year of publication, tumor type, ethnicity, number of cases and controls, minor allele frequency (MAF) in controls, genotype frequency for cases and controls, characteristics for cancer cases, genotyping methods.

Statistics

The odds ratios (ORs) with 95% confidence intervals (95% CIs) of cancer associated with the MLH1 polymorphisms were estimated for each study. For each polymorphism, we evaluated the risk of the variant homozygous compared with the wild-type homozygous (AA vs. GG for -93G>A and VV vs. II for I219V, respectively). Then we calculated the ORs with both dominant and recessive genetic models of the variant allele for the two polymorphisms. Subgroup analyses, according to tumor type (if a tumor type contains less than two individual studies, it was combined into the “other tumors” group), ethnicity (categorized as non-Asian and Asian, as the “mixed ethnicity” of four studies were composed of individuals not from Asia), family history, age at diagnose were also conducted. For each subgroup, statistical heterogeneity between studies was assessed using the Chi-square-based Q test. The heterogeneity was considered significant when I^2>50% (Lau et al., 1997). The fixed-effects model based on the Mantel-Haenszel method and the random-effects model based on the DerSimonian method were use to combine values from each studies. These two models provide similar results when the group of studies was considered homogeneity; otherwise, the random effect model is more appropriate. We also performed cumulative meta-analysis to evaluate whether the summary estimates for the allele contrasts changed over time as more data accumulated (Lau et al., 1992). An estimate of potential publication and other selection bias was investigated by the inverted funnel plots and the Egger’s test (Egger et al., 1997). An asymmetric plot indicates a possible publication bias. In the Egger’s linear regression test, P<0.05 was considered statistically significant publication bias.

All analyses were done in the STATA software (version 10.1, Stata Corporation, College Station, TX) and Review Manage (version 5.0.25, The Cochrane Collaboration, Oxford, United Kingdom, 2010). All the P values were two-sided.

Results

Characteristics of Studies

There are 20 case-control studies concerning -93G>A polymorphism (Ito et al., 1999; Deng et al., 2003; Park et al., 2004; Chen et al., 2005; Lee et al., 2005; Beiner et al., 2006; Song et al., 2006; Raptis et al., 2007; An et al., 2008; Harley et al., 2008; Koessler et al., 2008; Scott et al., 2008; Tulupova et al., 2008; Campbell et al., 2009; Shi et al., 2010; Shih et al., 2010; van Roon et al., 2010; Lacey et al., 2011; Lo et al., 2011; Whiffin et al., 2011) and 18 studies for I219V (Mathonnet et al., 2003; Listgarten et al., 2004; Lee et al., 2005; Mei et al., 2006; Song et al., 2006; Raptis et al., 2007; An et al., 2008; Capella et al., 2008; Smith et al., 2008; Campbell et al., 2009; Conde et al., 2009; Joshi et al., 2009; Nejda et al., 2009; Tanaka et al., 2009; Langeberg et al., 2010; Picelli et al., 2010). The characteristics including tumor type, the ethnicity of population, number of cases and controls for each MLH1...
polymorphisms are summarized in Table 1. Most studies indicated that the genotypes distribution in the controls was consistent with Hardy-Weinberg equilibrium except for six studies for -93G>A (Park et al., 2004; Beiner et al., 2006; An et al., 2008; Harley et al., 2008; Scott et al., 2008; Shih et al., 2010) and four studies for I219V (Listgarten et al., 2004; An et al., 2008; Conde et al., 2009; Nejda et al., 2009). Diverse genotyping methods were used, including polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP), TaqMan, and matrix-assisted laser desorption ionization time-of-flight (Chip-based MALDI-TOF) mass spectrometry.

Quantitative Synthesis

MLH1 -93G>A: The eligible studies included 24137 cancer cases and 21530 controls. There was significant difference for the variant A allele frequency between the two groups (Asian, OR = 0.52; 95% CI: 0.46-0.58; non-Asian, OR = 0.21; 95% CI: 0.20-0.22; P < 0.001; Figure 1). The evaluations of the association between MLH1 -93G>A with cancer risk are presented in Table 2. Overall, significant increased risk was detected in codominant model (AA vs. GG, OR = 1.22; 95% CI: 1.03-1.44), as well as in dominant and recessive model (AA+AG vs. GG, OR = 1.10; 95% CI: 1.00-1.22; AA vs. AG+GG, OR = 1.20; 95% CI: 1.06-1.36, respectively; Figure 2). The increased risks were more predominant in the non-Asian individuals carrying the A allele compared with those with the G allele in all genetic models (AA vs. GG, OR = 1.28; 95% CI: 1.04-1.58; AA+AG vs. GG; OR = 1.15; 95% CI: 1.04-1.28; AA vs. AG+GG; OR = 1.45; 95% CI: 1.11-1.89). However, no evaluated risk was found in Asian group. In subgroup analyses, an increased frequency of the -93A genotype was observed in all models tested (AA vs. GG, OR = 1.20; 95% CI: 1.11-1.89, respectively; Figure 2). The increased risks were more predominant in the non-Asian individuals carrying the A allele compared with those with the G allele in all genetic models (AA vs. GG, OR = 1.28; 95% CI: 1.04-1.58; AA+AG vs. GG; OR = 1.15; 95% CI: 1.04-1.28; AA vs. AG+GG; OR = 1.45; 95% CI: 1.11-1.89). However, no evaluated risk was found in Asian group. In subgroup analyses, an increased frequency of the -93A genotype was observed in all models tested among individuals with a positive family history and an age>50 when diagnosed (AA vs. GG; OR = 1.94; 95% CI: 1.18-3.19; OR = 1.71, 95% CI: 1.06-2.75, respectively). In the analyses stratified by tumor type, the variant genotype of -93G>A was a risk factor of colorectal cancer in the
Table 2. MLH1 -93G>A Polymorphism and Cancer Risk

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Ca/Co</th>
<th>OR (95% CI)</th>
<th>I²</th>
<th>OR (95% CI)</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>20</td>
<td>24137/21530</td>
<td>72</td>
<td>1.10 (1.00-1.22)</td>
<td>77</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>9</td>
<td>3369/3221</td>
<td>75</td>
<td>0.98 (0.74-1.29)</td>
<td>78</td>
</tr>
<tr>
<td>Non-Asian</td>
<td>11</td>
<td>20768/18309</td>
<td>71</td>
<td>1.15 (1.04-1.28)</td>
<td>78</td>
</tr>
<tr>
<td><strong>Tumor types</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung cancer</td>
<td>4</td>
<td>533/374</td>
<td>70</td>
<td>1.52 (0.98-2.35)</td>
<td>76</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>5</td>
<td>2294/3934</td>
<td>72</td>
<td>1.35 (1.03-1.77)</td>
<td>74</td>
</tr>
<tr>
<td>Other cancers</td>
<td>3</td>
<td>268/2812</td>
<td>0</td>
<td>1.47 (1.13-1.90)</td>
<td>40</td>
</tr>
<tr>
<td>Family history+</td>
<td>3</td>
<td>2347/2812</td>
<td>69</td>
<td>1.33 (1.10-1.61)</td>
<td>60</td>
</tr>
<tr>
<td><strong>Tumor sizes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at diagnose &lt;50</td>
<td>4</td>
<td>1756/1909</td>
<td>82</td>
<td>0.98 (0.63-1.53)</td>
<td>86</td>
</tr>
<tr>
<td>Age at diagnose &gt;50</td>
<td>5</td>
<td>1068/1168</td>
<td>40</td>
<td>1.25 (0.76-2.07)</td>
<td>88</td>
</tr>
<tr>
<td>Family history+</td>
<td>3</td>
<td>439/1068</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>
| **OR, odds ratio; CI, confidence interval; I² of Q-test for heterogeneity test; A random-effect model was used when I²>50%**

Table 3. MLH1 I219V Polymorphism and Cancer Risk

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Ca/Co</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>18</td>
<td>10632/11231</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5</td>
<td>1575/1649</td>
</tr>
<tr>
<td>Non-Asian</td>
<td>17</td>
<td>9155/9912</td>
</tr>
<tr>
<td><strong>Tumor types</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>7</td>
<td>5449/4710</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>4</td>
<td>1402/1636</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>2</td>
<td>1428/1367</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>2</td>
<td>791/813</td>
</tr>
<tr>
<td>Other cancer</td>
<td>3</td>
<td>1553/2705</td>
</tr>
</tbody>
</table>

**Test of Heterogeneity**

Substantial heterogeneity was observed among the 20 studies that included the -93G>A polymorphism but not among the 18 studies that include the I219V polymorphism. To account for heterogeneity, we conducted several subgroup meta-analyses. Significant heterogeneity was detected between different tumor types and ethnicities. We evaluated the source of heterogeneity for all the genetic models by tumor type, ethnicity and sample size. We found that none of the tumor type, ethnicity and sample size (dominant model: Chi² = 1.21, df = 4, P = 0.88; Chi² = 1.14, df = 1, P = 0.29; Chi² = 0.00, df = 1, P = 0.96; respectively) contributed to substantial altered heterogeneity, which was consistent with further meta-regression (dominant model: P = 0.698, P = 0.196, P = 0.932, respectively) that none of the three could explain significant between-study heterogeneity.

**Publication Bias**

For -93G>A, the magnitude of the summary ORs has been undergoing a trend of an increasingly stable effect as postulated (in random effect model, summary OR for all the genetic models by tumor type, ethnicity and sample size. We found that none of the tumor type, ethnicity and sample size (dominant model: Chi² = 1.21, df = 4, P = 0.88; Chi² = 1.14, df = 1, P = 0.29; Chi² = 0.00, df = 1, P = 0.96; respectively) contributed to substantial altered heterogeneity, which was consistent with further meta-regression (dominant model: P = 0.698, P = 0.196, P = 0.932, respectively) that none of the three could explain significant between-study heterogeneity.
Discussion

On the basis of 33 case-control studies focused on the two polymorphisms -93G>A and I219V of MLH1, our meta-analysis provided evidence that variant homozygote AA of -93G>A was associated with a modest but significantly increased risk of cancers, especially among non-Asian individuals. Significant associations emerged in the subgroup analysis of endometrial cancer.

No association was observed in current analysis between I219V and cancer risk. This is consistent with several functional analyses indicated that the variant has binding properties (to PMS2) and DNA repair efficiency similar to the wild type (Kondo et al., 2003; Raevaara et al., 2005). However, Kim et al. (2004) suggested homozygosity for the 219V variant was correlated with reduced MLH1 expression significantly among sporadic colorectal cancer cases from Korea. Further and larger samples study need to identify the association and the role of polymorphism I219V in carcinogenesis.

Epigenetic silencing of the MLH1 gene is one of the most common causes of DNA mismatch repair capacity deficiency (Goodfellow et al., 2003). Some study has indicated that the SNP -93G>A may increase the susceptibility of the promoter sequence to methylation. Because this polymorphism is located in a CpG island (which occurs at the cytosines of the CpG dinucleotides, and often occurs in clusters), adjacent to CpG cites that are able to undergo methylation (Deng et al., 2001). Goodfellow et al. (2003) reported >20% MLH1 gene promoter methylation occurred in a large series of endometrial carcinomas. Chen et al. (2007) have showed the variant alleles of SNP -93G>A on MLH1 was associated with MLH1 gene methylation in endometrial and colorectal cancers. They further confirmed that the -93 SNP may affect MLH1 gene expression by altering protein binding to the promoter of MLH1 gene. This was consistent with the conclusions of Mrkonjic et al. (2010). As the MLH1 promoter is bi-directional with the EPM2AIP1 gene located on the antisense strand, Perera et al. (2011) showed that the -93G>A polymorphism modifies the efficiency both of MLH1 and EPM2AIP1 transcription. EPM2AIP1 encodes a protein binding to laforin, also its function is unknown. Consistent with these observations, our meta-analysis showed that individuals carrying the AA genotype were associated with a higher risk of cancer than subjects with the AG+GG genotype.

Among older individuals (>50y) or subjects with a cancer family history, significant associations were also observed in all genetic models. This has shield light on that different genetic background may modify the cancer risk although relative fewer studies included in our meta-analysis. There appeared to be ethnicity-specific genetic effects. A clear association between MLH1 -93G>A and cancer risk was indicated in non-Asian individuals, suggesting genetic diversity among different ethnicities.

Several concerns may account for it. First, carcinogenesis is a multifactor progress with different incidence in different populations. As the average MAF was 0.52 in Asian group, which was higher than that in the non-Asian group (0.21). The difference in MAF might be a reflection of environmental impact on gene distribution (Hirschhorn et al., 2002). Environmental exposures, such as alcohol consumption, exposure of exogenous carcinogens, different life style between Europeans and Asians, which are corrected in part by the involvement of carcinogens, different life style between Europeans and Asians, which are corrected in part by the involvement of carcinogens, DNA repair efficiency etc. Differences in genetic background may modify the cancer risk.
gene-to-gene and gene-to-environment interactions. Second, the influence of MLH1 polymorphisms on cancer risk may be overwhelmed by the presence of other unknown genes which may involve in the cancer development. Also, other potential factors such as selection bias, different matching criteria and misclassifications on disease status and genotyping may also be involved.

There are some limitations in this meta-analysis. First, our results were unadjusted, while a more precise analysis needs to be executed with individual data such as age, sex and smoking pack-years. Second, the sample size for Asians was much smaller than for non-Asians. Thus, the analysis in Asian group may be underpowered. Large sample studies should further address the associations between MLH1 -93G>A and cancer risk in Asians. Although the limitations listed above may lead to confounding bias, advantages in our meta-analysis should also be acknowledged. Studies included in the current meta-analysis strictly and satisfactory met our selection criteria. In addition, a well designed systematic review of publication studies on the association of MLH1 polymorphisms and cancer risk is statistically more powerful than any single study.

In summary, consistent with functional evaluations, our study supports that MLH1 polymorphism -93G>A, but not I219V, may contribute to individual susceptibility of cancers. Large sample association studies and studies assessing gene-to-gene as well as gene-to-environment interactions are required to confirm these findings.

Acknowledgements

This work was supported by National Natural Science Funds (Grant Nos. 81071643 and 30772549).

References


MLH1 Polymorphisms and Cancer risk: a Meta-analysis Based on 33 Case-control Studies

DOI:http://dx.doi.org/10.7314/APJCP2012.13.3.901


