A polymorphism in the hMLH1 gene (-93G→A) associated with lung cancer susceptibility and prognosis

CHUEN-MING SHIH¹, CHIH-YI CHEN², I-HSUAN LEE³, WEI-TING KAO³ and YI-CHING WANG³

¹Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine; ²Cancer Center, China Medical University Hospital, Taichung; ³Department of Pharmacology, College of Medicine, National Cheng Kung University, Tainan, Taiwan, R.O.C.

Received July 15, 2009; Accepted September 21, 2009

DOI: 10.3892/ijmm_00000327

Abstract. Polymorphisms in DNA repair genes may be associated with differences in repair capacity of DNA damage and may thereby influence an individual’s susceptibility to lung cancer. We investigated the association between the -93G→A polymorphism in the mismatch repair hMLH1 gene for its role in the susceptibility and survival of non-small cell lung cancer (NSCLC) patients. Using a case-control study design, 165 NSCLC patients and 193 controls with similar range for age, gender and smoking habit distributions were subjected to genotype analysis. The risk of lung cancer was estimated by logistic regression analysis. The Kaplan-Meier method was used to estimate the probability of survival and the log-rank test was used to assess the significance of the difference between survival probabilities. The homozygous variant A/A genotype was associated with a significantly increased risk for lung cancer compared with other patients. The other genotypes (Crude analysis P=0.003, Adjusted analysis P=0.011, using the logistic regression model). The patients with a homozygous variant A/A genotype had a trend toward poorer prognoses compared with other patients, and smoking habit distributions were subjected to genotype analysis. The risk of lung cancer was estimated by logistic regression analysis. The Kaplan-Meier method was used to estimate the probability of survival and the log-rank test was used to assess the significance of the difference between survival probabilities. The homozygous variant A/A genotype was associated with a significantly increased risk for lung cancer compared with other patients (Crude analysis P=0.003, Adjusted analysis P=0.011, using the logistic regression model). The patients with a homozygous variant A/A genotype had a trend toward poorer prognoses compared with other patients, especially smoking (P=0.05, by log-rank test), male (P=0.06), or squamous carcinoma (P=0.08) patients. This is the first case-control study to show a significant association between the hMLH1-93G→A polymorphism and the susceptibility to and prognosis of lung cancer. The results herein may be useful for risk assessment and disease monitoring of NSCLC.

Introduction

Lung cancer is one of the most common malignancies in the world (1) and is the leading cause of cancer deaths in industrial countries. The major risk factor for lung cancer is an excessive exposure to tobacco smoke. However, only ~11% of tobacco smokers ultimately develop lung cancer (2), suggesting that genetic factors may influence the risk for lung cancer among those who are exposed to carcinogens. After the effect of tobacco smoke was stratified, an ~2.5-fold risk was attributable to a family history of lung cancer (2). Therefore, it is rational to speculate that certain common genetic variants or polymorphisms may have an impact on lung cancer risk. This genetic susceptibility may result from inherited polymorphisms in genes involved in carcinogen metabolism and repair of DNA damage (3-6). Our previous study found that hMLH1 was the major altered mismatch repair (MMR) gene involved in non-small cell lung cancer (NSCLC) tumorigenesis and that promoter methylation was the predominant mechanism in hMLH1 and hMSH2 deregulation (7). Several polymorphisms have been identified in the hMLH1 gene (8-10). We evaluated the association of the -93G→A polymorphism with lung cancer, since this polymorphism is located in the putative consensus sequence for the binding of transcription factor AP-4 (nCnnCAGCTG from -102 to -93), possibly influencing the activity of the hMLH1 promoter (11). This -93G→A polymorphism was not correlated with the risk of gastric carcinoma and hereditary nonpolyposis colorectal cancer (HNPPC) (11,12). But it has been associated with the risk of squamous cell carcinoma (SQ) of the lung, hepatocellular carcinoma, and breast cancer (8,13). Also, smoking is associated with MMR in colorectal neoplasia and suggests that the risk increase with smoking may differ with the hMLH1-93G→A genotype (14).

However, the -93G→A polymorphism of hMLH1 in relation to both cancer susceptibility and prognostic effect has never been examined in the same series of cancer patients and controls with similar range of age, gender and smoking distributions. Therefore, this study was designed to investigate the role of this polymorphism in the susceptibility to and the prognostic value of NSCLC using a case-control study.

Materials and methods

Study subject. A total of 165 non-small cell lung cancer patients who were admitted to China Medical University Hospital, Taichung, Taiwan, between 2000 and 2006 after obtaining appropriate Institutional Review Board permission and informed consent from the patients, were included in this study. They were included as experimental cases because no...
chemotherapy was given to these patients. Of them, 87 patients had adenocarcinomas (AD), and 63 had squamous cell carcinomas (SQ). The histological determination, including tumor types and stages, was performed according to the WHO classification method and the TNM system, respectively. Information on the smoking history of the lung cancer patients was obtained from hospital records. Subjects were categorized as non-smokers (never smokers) and smokers (including regular smokers if they smoked one cigarette per day for 6 months or longer and occasional smokers). The follow-up of 130 patients was performed at 2-month intervals in the first year after surgery and at 3-month intervals thereafter at outpatient clinics or by routine phone calls. The end of the follow-up period was March 2009 for all patients. The mean follow-up period for all patients was 17.1 months (range 0.5-84 months). For the 50 patients who survived the follow-up period (censored patients), the mean follow-up time was 21.6 months. For the 80 patients who died during the follow-up period, the mean follow-up period was 13.5 months.

Controls were randomly selected from a pool of healthy volunteers who visited the general health checkup center of China Medical University Hospital during the same period. A detailed questionnaire was completed for each case and control by a trained interviewer. The questionnaire included information on the average number of cigarettes smoked daily and the number of years the subjects had been smoking. A total of 193 healthy individuals, who had no known medical illness or hereditary disorders with similar range of age, gender and smoking distributions to case population, were studied as control subjects. They were gender-, age (± 3 yr)- and smoking (± 5 pack-yr)-matched subjects. Prior to commencement, this study was approved by the Research Ethics Committee of China Medical University and informed consent was obtained from each participant.

**hMLH1 genotyping.** Genomic DNA was extracted from peripheral blood lymphocytes by proteinase K digestion and phenol/chloroform extraction. The hMLH1-93G→A genotypes were determined by a PCR restriction fragment length polymorphism assay. The PCR primers for the -93G→A polymorphism were 5'-GTAGCGGGCAGTAGCCGCT-3' and 5'-CCGCCGAATAACCCCTGCCA-3', which generates a 259 fragment. The PCR reactions were performed in a 20 μl reaction volume containing 200 ng of genomic DNA, 25 pmol of each primer, 0.2 mM dNTP, 1X PCR buffer (50 mM KCl and 10 mM Tris-HCl, pH 8.3), 1.5 mM MgCl₂ and 1 unit Taq polymerase (Takara Shuzo, Otsu, Shiga, Japan). The PCR profile consisted of an initial denaturation step of 95°C for 5 min followed by 36 cycles of 95°C for 30 sec, 58°C for 30 sec, 72°C for 30 sec and a final elongation step of 72°C for 10 min. The PCR products were digested overnight with 10 units of PvuII (New England Biolabs, Beverly, MA) at 37°C and then resolved on a 1.5% agarose gel. The wild-type (G) allele (i.e., -93G) yielded 2 bands (134 and 125 bp) and the polymorphic (A) allele (i.e., -93A) was determined by the presence of the uncut 259 bp band (indicative of the absence of the PvuII cutting site) (Fig. 1). For quality control, the genotyping analysis was performed with blinding to case/control status and repeated twice for all subjects. The results of genotyping were 100% concordant. To confirm the genotyping results, selected PCR-amplified DNA samples (n=2 each for the GG, GA and AA genotypes) were examined by DNA sequencing, and the results were also 100% concordant.

**Statistical analysis.** Differences in clinical data between the lung cancer patients and the control subjects were checked by Mann-Whitney U test. Hardy-Weinberg equilibrium was assessed using a goodness-of-fit Chi-square (χ²) test for biallelic markers. The Fisher exact test was used to compare genotype distributions among different ethnic groups. The Pearson’s χ² test was used to compare genotype distributions between various clinical and genetic alteration factors in lung cancer patients. Statistical modeling, using logistic regression, was used to calculate the relative risk (odds ratio, OR) of homozygous variant A/A genotype to the wild-type G/G genotype and heterozygous variant G/A genotype for the case-control study. ORs were expressed together with the 95% confidence interval (CI). Multivariate logistic regression analysis was adjusted for age, gender and smoking habit. Type III censoring was performed on subjects who were still alive at the end of the study. The Kaplan-Meier method was used to estimate the probability of survival as a function of time and median survival. The log-rank test was used to assess the significance of the difference between homozygous variant A/A genotype and G/G and G/A genotypes among pairs of survival probabilities. Significance was accepted at P<0.05.

**Results**

**Distribution of the hMLH1-93G→A polymorphism in the general Taiwanese population and in other populations.** Genomic DNA from 193 non-cancer controls, who had no known medical illness or hereditary disorders with similar range of age, gender and smoking distribution to case population, was first analyzed to determine the distribution of the hMLH1 polymorphism in the Taiwanese population. The frequencies of the hMLH1 genotypes G/G, G/A and A/A found in the non-cancer controls in Taiwan were 18.7, 58.6 and 22.8%, respectively. The genotype distribution fit the expectation under Hardy-Weinberg equilibrium (P=0.98). The variant A and wild-type G allele frequencies in the non-cancer controls in Taiwan were 0.52 and 0.48, respectively. Table I shows the comparison of the distribution of the hMLH1 genotype in our controls with the data reported previously for other study populations. Chi-square (χ²) analysis indicated no
significant difference in the genotype distributions of hMLH1 between Taiwanese, Chinese (12), Japanese (11) and Korean (8) populations.

The relationship between the hMLH1-93G→A polymorphism and lung cancer risk. Table II shows the distribution of the hMLH1 polymorphism by case/control status in 193 control subjects and 165 cancer patients and the clinico-pathological parameters of the lung cancer patients. Since hMLH1 is a tumor suppressor-like gene, further logistic regression analyses were assessed by a recessive-effect model, which considered the homozygous variant A/A as a risk genotype. Overall, there was a significant difference in genotype distributions between the non-cancer controls and the lung cancer patients. The homozygous variant A/A genotype was associated with a significantly increased risk for lung cancer compared with the combined G/G and G/A genotypes (Crude analysis P=0.003, Adjusted analysis P=0.011, using the logistic regression model). When the patient group was stratified by gender, tumor type, tumor stage and smoking history, an increased

Table I. Frequency of A alleles of the hMLH1 (G→93A) polymorphism in different populations.

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Population</th>
<th>G/G (%)</th>
<th>G/A (%)</th>
<th>A/A (%)</th>
<th>A allele frequency</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taiwanese</td>
<td>n=193</td>
<td>36 (18.7)</td>
<td>113 (58.6)</td>
<td>44 (22.8)</td>
<td>0.52</td>
<td>The present study</td>
</tr>
<tr>
<td>Chinese</td>
<td>n=54</td>
<td>8 (14.9)</td>
<td>27 (50.0)</td>
<td>19 (35.2)</td>
<td>0.57</td>
<td>0.097</td>
</tr>
<tr>
<td>Korean</td>
<td>n=371</td>
<td>71 (19.2)</td>
<td>206 (55.5)</td>
<td>94 (25.3)</td>
<td>0.53</td>
<td>0.511</td>
</tr>
<tr>
<td>Japanese</td>
<td>n=84</td>
<td>22 (26.2)</td>
<td>46 (54.8)</td>
<td>16 (19.0)</td>
<td>0.46</td>
<td>0.247</td>
</tr>
</tbody>
</table>

*aP-values were calculated using the Fisher exact test.

Table II. Distribution of the hMLH1 (G→93A) polymorphism by case/control status and clinico-pathological parameters of lung cancer patients.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G/G (%)</th>
<th>G/A (%)</th>
<th>A/A (%)</th>
<th>Total</th>
<th>Crude OR (95% CI)*</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-cancer control</td>
<td>36 (18.7)</td>
<td>113 (58.6)</td>
<td>44 (22.0)</td>
<td>193</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>41 (24.9)</td>
<td>64 (38.8)</td>
<td>60 (36.4)</td>
<td>165</td>
<td>1.94 (1.22-3.07)*</td>
<td>1.84 (1.15-2.93)</td>
</tr>
<tr>
<td>Gender</td>
<td>29 (22.8)</td>
<td>53 (41.7)</td>
<td>45 (35.4)</td>
<td>127</td>
<td>1.86 (1.13-3.05)*</td>
<td>1.52 (0.90-2.58)</td>
</tr>
<tr>
<td>Smoking habit</td>
<td>12 (31.6)</td>
<td>11 (29.0)</td>
<td>15 (39.5)</td>
<td>38</td>
<td>2.21 (1.06-4.59)*</td>
<td>7.80 (1.98-30.7)</td>
</tr>
<tr>
<td>Age ≥65</td>
<td>24 (25.3)</td>
<td>36 (37.9)</td>
<td>35 (36.9)</td>
<td>95</td>
<td>1.98 (1.16-3.38)*</td>
<td>2.29 (1.29-4.05)</td>
</tr>
<tr>
<td>Age &lt;65</td>
<td>17 (24.3)</td>
<td>28 (40.0)</td>
<td>25 (35.7)</td>
<td>70</td>
<td>1.88 (1.04-3.41)*</td>
<td>1.55 (0.78-3.09)</td>
</tr>
<tr>
<td>Tumor type</td>
<td>19 (21.8)</td>
<td>34 (39.1)</td>
<td>34 (30.1)</td>
<td>87</td>
<td>2.17 (1.29-3.75)*</td>
<td>2.22 (1.28-3.90)</td>
</tr>
<tr>
<td>Tumor stage</td>
<td>15 (23.8)</td>
<td>28 (44.4)</td>
<td>20 (31.8)</td>
<td>63</td>
<td>1.58 (0.84-2.95)</td>
<td>1.28 (0.67-2.44)</td>
</tr>
</tbody>
</table>

*aOdds ratios (ORs) were calculated to measure the association of the A/A genotype with lung cancer risk, with that of the G/G+G/A genotypes being referred to as 1. CI, confidence interval. P-values with significance were labeled in superscript. *Adjusted for age, gender and smoking; for gender, adjusted for age and smoking; for age, adjusted for gender and smoking; for smoking, adjusted for gender and age. AD, adeno-carcinomas; SQ, squamous cell carcinomas.
The frequency of the A/A genotype was observed in female (adjusted OR=7.80; 95% CI: 1.98-30.70; P=0.003), non-smoker (P=0.003) and AD (P=0.006) patients in adjusted logistic regression model. In addition, a significant difference between the advanced patients (stages III and IV) patients and the controls (adjusted OR=2.13; 95% CI: 1.28-3.55; P=0.004) was observed.

The prognostic significance of the hMLH1-93→A polymorphism in non-small cell lung cancer. All patients were followed after surgery. However, 35 patients were lost to follow-up during our study period. After several training tests, we found that the homozygous variant A/A genotype had the strongest prognostic effect. There was a trend toward a shorter survival in those patients with the A/A genotype (P=0.19, by log-rank test) compared with those with the G/A and G/G genotypes, among all 130 patients who had survival data available. The prognostic trends of lung cancer related to the A/A genotype of the hMLH1 polymorphism were also confirmed in smoker (P=0.05), male (P=0.06) and SQ patients (P=0.08) (Fig. 2).

Discussion

To our knowledge, this is a case-control study with the largest sample number of hMLH1-93G→A polymorphisms in relation to lung cancer susceptibility and prognosis. In this study, only patients who had undergone surgery without chemotherapy were enrolled. These patients had a more accurate pathological stage than the patients who did not receive surgery and the prognosis analysis would not be complicated by the chemotherapy. Our results indicated that the effect of the hMLH1-93A/A variant genotype on the risk of lung cancer was more evident in non-smoking female AD patients, whereas on the prognosis was more apparent in smoking male SQ patients.

Since hMLH1 is a tumor suppressor-like gene, further logistic regression analyses were assessed by a recessive-effect model, which considered the homozygous variant A/A as a risk genotype. We also performed a trend test by a dominant-effect model, which combined A/A and G/A variant genotypes to compare with the wild-type G/G genotype. The data indicated an increased risk for lung cancer with A/A or G/A variant genotype compared with the wild-type G/G genotype (data not shown). The presence of an underlying effect of hMLH1-93G→A polymorphism, perhaps results in tumorigenesis because of haplo-insufficiency merits further clarification. The wild-type T/T and perhaps heterozygous G/A polymorphisms may confer an efficient DNA repair in the tumor. Therefore, the hMLH1-93A/A variant genotype may be involved in lung cancer susceptibility by encoding a defective protein functioning in the repair of DNA damage.

Figure 2. The Kaplan-Meier survival curves with respect to the hMLH1-93G→A polymorphism in relation to clinicopathological parameters of lung cancer patients. (A) In all patients, (B) in male patients, (C) in smoker patients and (D) in patients with squamous cell carcinoma (SQ). The P-value for each analysis is indicated.
and replication error correction due to their relative low level compared to the wild-type genotype. The homozygous variant A/A genotype may encode the lowest hMLH1 protein level, thereby leading to the worst prognosis for the patients. We propose that the hMLH1-93G→A polymorphism may change the protein level or biological function of hMLH1 protein and may be involved in carcinogenesis to some extent.

The mechanism responsible for the association between the hMLH1-93G→A polymorphism and NSCLC remains to be elucidated. Mismatch repair enzymes recognize and repair mismatched DNA base pairs that occur during replication (15,16). Polymorphisms in hMLH1 may play a role in NSCLC (7,17-19). The hMLH1-93G→A polymorphism involves a substitution of G→A in the promoter region of hMLH1. Although one in vitro study found no differences between the G and A alleles in promoter activity (12), the polymorphism is located within a putative nuclear factor for the interleukin-6 expression (NF-IL6) binding site and may alter hMLH1 expression. In addition, this polymorphism is located in the putative consensus sequence for the binding of transcription factor AP-4 (-102 to -93), possibly influencing the activity of the hMLH1 promoter (11). Further functional assays will provide the ultimate answer for the expression influence of this polymorphism.

The result suggested that the A/A genotype was associated with a significantly increased risk for AD lung cancer. However, an increased risk of this genotype with SQ lung cancer was observed in a Korean population (8). The discrepancies may result from recruitment bias, heterogeneity of cohort analyzed and histology-dependent association in studies from different geographic areas. The present study showed that the association of the hMLH1-93G→A polymorphism may be a risk factor for non-smoking female AD lung cancer in Taiwan (Table II). Registry data indicate a low male-to-female ratio of 2:1 for lung cancer mortality in Taiwan (20). However, few Taiwanese females smoke cigarettes. The distribution of cigarette smokers in Taiwan is 59.4% for males and 3.8% for females (20). If hMLH1-93G→A polymorphism is a susceptible genotype, its high prevalence in female lung cancer patients may partly explain their high rate of AD lung cancer in Taiwan.

Our study demonstrated that the homozygous variant A/A hMLH1 genotype is a prognostic factor for decreased survival in NSCLC especially for smoking male SQ patients (Fig. 2). Tumor stage was not a confounding factor of these analyses because smoking habits and tumor types were not associated with tumor stage in this cohort of patients (data not shown). Several studies also indicated that neoplasms with deficient mismatch repair have a significantly worse prognosis compared to those with intact mismatch repair (7). It is possible that the hMLH1 variant A/A genotype confers a low level of repair and correlates with worse survival in smoking NSCLC patients whose genome contains more DNA damage-induced mismatches, suggesting that the effect of smoking may work corporately with the genetic predisposition. It also seems possible that the hMLH1-93G→A polymorphism may be a genetic marker of other genes that affect the susceptibility and prognosis of lung cancer patients.

In conclusion, the present study found a significant association between the hMLH1-93G→A polymorphism and the susceptibility to non-smoking female AD lung cancer. Also, the hMLH1-93G→A polymorphism was associated with the prognosis of smoking male SQ lung cancer. However, given the number of comparisons and the sample size of the current study, the conclusions should be interpreted with caution and confirmed by other ethnic populations. It should be noted that little is know about the hMLH1-93G→A polymorphism in terms of the potential impact on cancer risk and prognosis of Caucasian or other populations. The use of the hMLH1-93G→A polymorphism may supplement current clinical evaluation methods for risk assessment in population studies and possibly for disease monitoring of lung cancer.

Acknowledgements

This study was supported in part by grant NSC96-2628-B-006-048-MY3 from the National Science Council and grant DOH97-TD-G-111-035 from the Department of Health (The Executive Yuan, Republic of China).

References