Functional haplotypes in the promoter of matrix metalloproteinase-2 and lung cancer susceptibility

Yifeng Zhou1, Chunyuan Yu1, Xiaoping Miao1, Yonggang Wang2, Wen Tan1, Tong Sun1, Xuelei Zhang1, Ping Xiong1,3 and Dongxin Lin1

1Departments of Etiology and Carcinogenesis and 2Department of Thoracic Surgery, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
3Current address: Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

Matrix metalloproteinase-2 (MMP-2) plays important roles in cancer initiation and progression. Our previous studies revealed that the −1306C→T and −735C→T polymorphisms in MMP2 promoter significantly influence transcriptional activity and their genotypes and haplotypes are associated with susceptibility to several cancers. This case–control study examined the contribution of these two polymorphisms to the risk of developing lung cancer. MMP2 genotypes and haplotypes were determined in 770 cases and 777 controls and the associations with risk of lung cancer were estimated by logistic regression. We observed a 2-fold [odds ratio (OR), 2.12; 95% confidence interval (CI), 1.64–2.72] or 1.6-fold (OR, 1.57; 95% CI, 1.27–1.95) excess risk of developing lung cancer for the −1306CC or −735CC genotype carriers compared with non-carriers, respectively. A greater risk of lung cancer was associated with the C−1306C–T−735 haplotype (OR, 5.01; 95% CI, 2.57–9.78) compared with the T−1306C–T−735 haplotype, suggesting a synergistic effect of these two polymorphisms. Furthermore, a greater than additive joint effect of the polymorphisms and smoking increased an even higher risk of lung cancer. The OR for smokers with the C−1306C–T−735 haplotype was 6.24 (95% CI, 4.51–8.64), which was significantly higher than that (OR, 4.10; 95% CI, 2.89–5.81) of smokers with the T−1306C or T−735− allele containing haplotypes (P < 0.001). These results are consistent with our previous findings and further support the hypothesis that gain-of-function of MMP2 resulting from genetic polymorphisms plays an important role in human carcinogenesis.

Introduction

Matrix metalloproteinase-2 (MMP-2), among other MMPs, plays a central role in degrading or breaking down extracellular matrix, a physical barrier in restraining the expanding growth and migration of tumor cells (1–3). Increased expression of MMP-2 has been associated with cancer invasive and metastatic phenotypes (4–6). However, expanding studies have demonstrated that, in addition to destroying extracellular matrix to permit tumor cell invasion and metastasis, MMP-2 also plays important roles in several steps of cancer development (7,8). New specific matrix and non-matrix substrates for MMP-2 have been identified and the biological consequences of cleavage of these substrates indicates that MMP-2 is involved in cell proliferation, apoptosis, angiogenesis and immune surveillance (7). For instance, MMP-2 can cleave insulin-like growth factors (IGFs)-binding proteins and release IGFs (9), which are well known to have a strong effect on stimulating cell proliferation and inhibiting apoptosis. Overexpression of MMP-2 has also been detected in precancerous lesions and early stage of human cancer (10,11), suggesting that it may be an early event in carcinogenesis. Furthermore, it has been demonstrated that high levels of constitutive expression of MMP-2 is associated with a high susceptibility to tumor formation in genetically modified animal models (12,13).

Human MMP2 gene exhibits sequence variations and several functional single nucleotide polymorphisms (SNPs) in the MMP2 promoter have been described (14–16). A C→T polymorphism located at nucleotide −1306 site disrupts an Sp1 regulatory element and the T allele thus, has a strikingly lower promoter activity compared with the C allele (14). We have recently shown that another C→T polymorphism located at nucleotide −735 also destroys an Sp1 binding element, with the T allele being associated with significantly diminished promoter activity (16). Interestingly, the −1306C→T and −735C→T polymorphisms are in a linkage disequilibrium and the T−1306C–T−735 haplotype displays an even lower promoter activity and mRNA expression compared with the haplotype consisting of only one T allele at the −1306 or −735 site, indicating an interactive effect of these two SNPs on MMP2 transcriptional function (16).

The association between the −1306C→T polymorphism and susceptibility to human cancers has been investigated in several studies. It has been shown that the −1306C allele is associated with an increased risk of common cancers, including lung, gastric cardia, breast, oral and colorectal cancers (17–21). Recently, we have explored the relationship between the −735C→T polymorphism, alone or in combination with the −1306C→T polymorphism, and risk of esophageal squamous cell carcinoma (SCC). We observed an increased risk of the cancer associated not only with the −1306C allele but also with the −735C allele; furthermore, an even greater association was observed between elevated risk of developing esophageal SCC and C−1306C– and C−735C-allele containing haplotypes (16).

In view of the important role that MMP-2 plays in cancer initiation and development, a more comprehensive study is warranted to evaluate the effect of MMP2 promoter haplotypes as a genetic modifier in the etiology of lung cancer, whose

Abbreviations: CI, confidence interval; MMP-2, matrix metalloproteinase-2; OR, odds ratio; SCC, squamous cell carcinoma; SNP, single nucleotide polymorphism.

1The first two authors contributed equally to this work.
rates of incidence and mortality have been increasing significantly and constantly in China and most other parts of the world. Here, we report a large contribution of MMP2 C–1306– and C–735–alleles containing haplotypes, alone or in combination with smoking, to the risk of developing lung cancer in a case–control study.

Materials and methods

Study subjects

Characteristics of the study subjects have been described previously (17). All subjects were ethnic Han Chinese. Patients were consecutively recruited between January 1997 and November 2001, at the Cancer Hospital, Chinese Academy of Medical Sciences (Beijing). All patients with histopathologically confirmed lung cancer were enrolled, yielding a 93% response rate; they were from Beijing City and surrounding regions. Controls were cancer-free individuals randomly selected from a nutritional survey consisting of 2500 individuals, which was conducted in the same period as the cases were collected. The selection criteria included no individual history of cancer and frequency matched to cases on sex and age (±5 years). Of the 781 cases and 852 controls who participated in the previous study (17), only 770 cases and 777 controls were successfully genotyped in this study because DNA samples of the rest of the subjects were no longer available. At recruitment, informed consent was obtained from each subject and each participant was then interviewed to obtain detailed information on demographic characteristics and lifetime history of tobacco use. This study was approved by the Institutional Review Board of the Chinese Academy of Medical Sciences Cancer Institute.

MMP2 genotyping

Genotypes at the −1306 and −735 site were determined by PCR-based denaturing high performance liquid chromatography and PCR-based restriction fragment length polymorphism methods as described previously (16,17). Genotyping was performed with blinding to case/control status; a 15% masked, random sample of cases and controls was tested twice by different persons; and the results were concordant for all masked duplicated sets.

Statistical analysis

The χ² test was used to compare the distributions of MMP2 genotypes and haplotypes between cases and controls. Smokers were categorized as light smokers and heavy smokers by the approximate 50th percentile pack year value among controls, i.e., <26 or ≥26 pack years [cigarettes per day/20] × (years smoked)]. Hardy–Weinberg equilibrium was tested by a goodness-of-fit χ² test. Linkage disequilibrium coefficient and haplotype frequencies were estimated using EH (EH-plus) and PHASE software, respectively (22,23). Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by unconditional logistic regression as a measure of association with risk of the development of lung cancer. The ORs were adjusted for age, sex and smoking status or pack years, where it was appropriate. A 2 × 2 test was used to compare the distributions of MMP2 genotypes and haplotypes among controls.

Table I. Distributions of select characteristics by case-control status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cases (n = 770)</th>
<th>Controls (n = 777)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (%)</td>
<td>Male</td>
<td>544 (70.6)</td>
<td>558 (71.8)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>226 (29.4)</td>
<td>219 (28.2)</td>
</tr>
<tr>
<td>Mean age in year (SD)</td>
<td>58.3 (10.9)</td>
<td>57.6 (7.6)</td>
<td>0.170</td>
</tr>
<tr>
<td>Smoking status, (%)</td>
<td>Non-smokers*</td>
<td>268 (34.8)</td>
<td>409 (52.6)</td>
</tr>
<tr>
<td></td>
<td>Smokers</td>
<td>502 (65.2)</td>
<td>368 (47.4)</td>
</tr>
<tr>
<td></td>
<td>&lt;26 pack years</td>
<td>169 (33.7)</td>
<td>184 (50.0)</td>
</tr>
<tr>
<td></td>
<td>≥26 pack years</td>
<td>333 (66.3)</td>
<td>184 (50.0)</td>
</tr>
<tr>
<td>Histological type (%)</td>
<td>SCC</td>
<td>318 (41.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>210 (27.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Otherb</td>
<td>242 (31.4)</td>
<td></td>
</tr>
</tbody>
</table>

*aNon-smokers were defined as subjects who smoked <10 cigarettes in a lifetime.

Materials and methods

Study subjects

Characteristics of the study subjects have been described previously (17). All subjects were ethnic Han Chinese. Patients were consecutively recruited between January 1997 and November 2001, at the Cancer Hospital, Chinese Academy of Medical Sciences (Beijing). All patients with histopathologically confirmed lung cancer were enrolled, yielding a 93% response rate; they were from Beijing City and surrounding regions. Controls were cancer-free individuals randomly selected from a nutritional survey consisting of 2500 individuals, which was conducted in the same period as the cases were collected. The selection criteria included no individual history of cancer and frequency matched to cases on sex and age (±5 years). Of the 781 cases and 852 controls who participated in the previous study (17), only 770 cases and 777 controls were successfully genotyped in this study because DNA samples of the rest of the subjects were no longer available. At recruitment, informed consent was obtained from each subject and each participant was then interviewed to obtain detailed information on demographic characteristics and lifetime history of tobacco use. This study was approved by the Institutional Review Board of the Chinese Academy of Medical Sciences Cancer Institute.

MMP2 genotyping

Genotypes at the −1306 and −735 site were determined by PCR-based denaturing high performance liquid chromatography and PCR-based restriction fragment length polymorphism methods as described previously (16,17). Genotyping was performed with blinding to case/control status; a 15% masked, random sample of cases and controls was tested twice by different persons; and the results were concordant for all masked duplicated sets.

Statistical analysis

The χ² test was used to compare the distributions of MMP2 genotypes and haplotypes between cases and controls. Smokers were categorized as light smokers and heavy smokers by the approximate 50th percentile pack year value among controls, i.e., <26 or ≥26 pack years [cigarettes per day/20] × (years smoked)]. Hardy–Weinberg equilibrium was tested by a goodness-of-fit χ² test. Linkage disequilibrium coefficient and haplotype frequencies were estimated using EH (EH-plus) and PHASE software, respectively (22,23). Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by unconditional logistic regression as a measure of association with risk of the development of lung cancer. The ORs were adjusted for age, sex and smoking status or pack years, where it was appropriate. A P-value of <0.05 was used as the criterion of statistical significance, and all statistical tests were two-sided. We tested the null hypotheses of additive and multiplicative gene smoking interactions and evaluated departures from additive and multiplicative interaction models (24) by including main effect variables and their product terms in the logistic regression model. A more than additive interaction was indicated when OR11 > OR10 + OR01 − 1, where OR11 = OR when both factors are present, OR10 = OR when only factor 1 is present, and OR01 = OR when only factor 2 is present. A more than multiplicative interaction was suggested when OR11 > OR01 × OR10. All statistical analyses were performed with computer programs from Statistical Analysis System (version 6.12; SAS Institute, Cary, NC).

Results

The relevant characteristics of study subjects are shown in Table I. There were no statistically significant differences between the cases and controls in terms of age and sex distributions (P = 0.170 and 0.613, respectively). However, as expected, more smokers were presented among cases compared with controls (65.2% versus 47.4%; P < 0.001). Furthermore, the cases had higher values of pack years smoked than the controls; 66.3% of smokers among the cases smoked an equivalent of ≥26 pack years, whereas this value was only 50.0% among the controls (P < 0.001). Among the cases, 318 (41.5%) were classified as SCC, 210 (27.3%) as adenocarcinoma (AC) and 242 (31.4%) as other types, including undifferentiated cancer (n = 93), bronchoalveolar carcinoma (n = 85) and mixed cell carcinoma (n = 64).

Genotyping data are presented in Table II. The frequencies of three MMP2 −1306 or −735 genotypes among controls were not different from those expected from the Hardy–Weinberg equilibrium (P = 0.842 and 0.374). However, the genotype frequencies at both sites were significantly different between cases and controls (both P < 0.001), with the −1306CC and −735CC genotypes being overrepresented in cases. Multivariate analysis showed that the −1306CC genotype carriers had a 2-fold elevated risk for developing lung cancer (adjusted OR, 2.12; 95% CI, 1.64–2.72) compared with the non-carriers. Similarly, the −735CC genotype carriers were also at an increased risk of developing the cancer, compared with the TT or CT genotype carriers (adjusted OR, 1.57; 95% CI, 1.27–1.95). Moreover, the increased risk associated with the −735CC or −735CC genotype was consistently observed among different subtypes of lung cancer (Table II).

Linkage disequilibrium analysis showed that the −1306C→T and −735C→T polymorphisms were in a linkage disequilibrium (P < 0.001). The χ² test of statistical significance for a two-locus disequilibrium gave a test statistic value of 10.2 (D² = 0.53) for the cases, 11.0 (D² = 0.33) for the controls and 14.4 (D² = 0.33) for all subjects. MMP2 haplotypes in cases and controls were then constructed by using the PHASE software and the results are presented in Table III. The distribution of haplotype frequencies differed significantly between the cases and controls (χ² = 55.6, P < 0.0001, df = 3), with the T–1306 or T–735-containing haplotypes being more prevalent among controls than among cases. Compared with the T–1306–T–735 haplotype, each of the other haplotype containing at least one −1306C or −735C allele was associated with an increased risk of lung cancer; the adjusted ORs for the T–1306C–735, C–1306–T–735 and C–1306C–735 haplotypes were 3.01 (95% CI, 1.47–6.18), 3.81 (95% CI, 1.89–7.65) and 5.01 (95% CI, 2.57–9.78), respectively (trend test, P < 0.001).

The risk of lung cancer related to MMP2 genotypes and haplotypes were further examined with stratification by smoking status and pack year values (Tables IV and V). It was found that among non-smokers, the adjusted OR of lung cancer for...
Table II. Genotype and allele frequencies of MMP2 in cases and controls and their association with risk of lung cancer

<table>
<thead>
<tr>
<th>MMP2 genotype</th>
<th>Controls</th>
<th>Overall cases</th>
<th>Cases with SCC</th>
<th>Cases with AC</th>
<th>Cases with other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>−1306 C/T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>539 (69.4)</td>
<td>635 (82.5)</td>
<td>268 (84.3)</td>
<td>173 (82.4)</td>
<td>194 (80.2)</td>
</tr>
<tr>
<td>CT</td>
<td>220 (28.3)</td>
<td>124 (16.1)</td>
<td>46 (14.5)</td>
<td>34 (16.2)</td>
<td>44 (18.2)</td>
</tr>
<tr>
<td>TT</td>
<td>18 (2.3)</td>
<td>11 (1.4)</td>
<td>4 (1.2)</td>
<td>3 (1.4)</td>
<td>4 (1.6)</td>
</tr>
<tr>
<td>C allele frequency</td>
<td>0.84</td>
<td>0.91</td>
<td>0.92</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>Adjusted ORb (95% CI)</td>
<td>2.12 (1.64-2.72)</td>
<td>2.37 (1.67-3.37)</td>
<td>2.06 (1.38-3.10)</td>
<td>1.78 (1.24-2.58)</td>
<td></td>
</tr>
<tr>
<td>−735 C/T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>425 (54.7)</td>
<td>506 (65.7)</td>
<td>214 (67.3)</td>
<td>136 (64.9)</td>
<td>156 (64.5)</td>
</tr>
<tr>
<td>CT</td>
<td>313 (40.3)</td>
<td>230 (29.9)</td>
<td>90 (28.2)</td>
<td>63 (30.1)</td>
<td>77 (31.8)</td>
</tr>
<tr>
<td>TT</td>
<td>39 (5.0)</td>
<td>34 (4.4)</td>
<td>14 (4.5)</td>
<td>11 (5.0)</td>
<td>9 (3.7)</td>
</tr>
<tr>
<td>C allele frequency</td>
<td>0.75</td>
<td>0.81</td>
<td>0.81</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>Adjusted ORb (95% CI)</td>
<td>1.57 (1.27-1.95)</td>
<td>1.70 (1.28-2.26)</td>
<td>1.52 (1.10-2.12)</td>
<td>1.50 (1.10-2.05)</td>
<td></td>
</tr>
</tbody>
</table>

*SCC, squamous cell carcinoma; AC, adenocarcinoma; other includes undifferentiated cancer (n = 93), bronchioalveolar carcinoma (n = 85) and mixed cell carcinoma (n = 64).

Data were calculated by logistic regression, with the MMP2 variant genotypes (CT or TT) as the reference group and adjusted for age, sex and smoking status.

Table III. Risk estimates for extended MMP2 promoter haplotypes in cases and controls

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>No. of chromosomes (%)</th>
<th>Adjusted ORb (95% CI)</th>
<th>Cases (n = 770, allele = 1540)</th>
<th>Controls (n = 777, allele = 1554)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T−1306C−735</td>
<td>11 (0.7)</td>
<td>52 (3.3)</td>
<td>1.00 (reference)</td>
<td></td>
</tr>
<tr>
<td>T−1306C−735</td>
<td>135 (8.8)</td>
<td>204 (13.1)</td>
<td>3.01 (1.47-6.18)</td>
<td></td>
</tr>
<tr>
<td>C−1306C−735</td>
<td>287 (18.6)</td>
<td>339 (21.8)</td>
<td>3.81 (1.89-7.65)</td>
<td></td>
</tr>
<tr>
<td>C−1306C−735</td>
<td>1107 (71.9)</td>
<td>959 (61.7)</td>
<td>5.01 (2.57-9.78)</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for age, sex and pack years values.

Discussion

We have previously reported a significant association between the MMP2 −1306CC genotype and increased risk of developing lung cancer (17). Recently, we have also identified a novel polymorphic site (−735C→T) as an Sp1 binding site and found that the −735CC genotype was associated with an increased MMP-2 expression and increased risk of the occurrence and metastasis of esophageal cancer. Furthermore, the −735CC and −1306CC genotypes were in a linkage disequilibrium and functioned interactively in the context of haplotype (16). In the present study, we extended the findings to lung cancer, showing that not only the −1306CC genotype but also the −735CC genotype was a susceptibility factor and these two genotypes had a strong interaction within a haplotype to influence the risk of cancer; the C−1306C−732 haplotype was associated with a 5-fold increased risk of developing lung cancer compared with the T−1306C or T−732-containing haplotypes. Moreover, we observed a more than additive joint effect between the risk genotypes or haplotypes and smoking, which was associated with even greater risk of lung cancer. These extended findings further support our notion that gain-of-function of MMP2 resulting from genetic polymorphisms confers an individual susceptibility to certain common cancers (16-19).

Overexpression of MMP-2 has been demonstrated to be an early event in the development of many types of human cancers, including lung cancer (10,11,26,27). Furthermore, this proteinase in lung tumors is produced not only by cancer cells but also by normal stromal cells and endothelial cells (28-30), suggesting that the overexpression of MMP-2 is probably owing to transcriptional changes but not gene amplification or activating mutation. Indeed, we and others have demonstrated that the −1306 and −735 genotypes and haplotypes in the MMP2 promoter are associated with significantly different transcriptional activity of the gene (14,16). Although the differences in allele transcription resulting from polymorphisms in the MMP2 promoters are subtle compared with the overexpression that arises from the amplification of oncogenes, the increased level of MMP-2 over a lifetime might foster an increased susceptibility to tumorigenesis. Several studies with genetically modified animals are consistent with this notion. It was found that when induced by carcinogenic stimulus, wild-type mice having the Mmp2 or Mmp9 gene developed more tumors than mice that lack these genes (31). Cancer cells injected through the vein were found to be more capable of colonizing the lungs of wild-type mice than the lungs of...
MMP2-knockout mice (12). Conversely, transgenic mice that overexpress membrane type MMP-1, a known activator of pro-MMP-2, were at an increased susceptibility to mammary tumor formation and metastasis (32). These experimental findings strongly support our observation that germline polymorphisms in the MMP2 promoter that cause high expression of MMP-2 over a lifetime may constitute an increased susceptibility to lung cancer.

Functional polymorphisms in some other MMP genes have also been linked to varying susceptibility to certain human cancers. It has been shown that a single guanine insertion polymorphism in the promoter region of MMP1 (2G allele), which creates an Est binding site and enhances transcriptional activity, has been associated with an increased risk for developing cancer of the lung, ovary, colo-rectum and endometrium compared with the 1G allele (33–36). A single adenosine deletion polymorphism in the MMP3 promoter (5A allele) has been associated with a higher risk for breast cancer and esophageal cancer, compared with the 5A allele which has reduced transcriptional activity (37,38). These findings are parallel to our results reported here showing that genetic polymorphisms, which enhance MMP2 transcriptional activity, are associated with an increased susceptibility to lung cancer.

A joint effect between the MMP2 promoter polymorphisms and tobacco smoking was evident in the present study. We found that the −1306CC and −735CC genotypes or the C−1306C−735 haplotype were significantly associated with lung cancer risk in non-smokers; however, the risk was markedly elevated in smokers, particularly heavy smokers, suggesting an additive gene smoking interaction. Tobacco smoking is an established etiological factor for lung cancer; therefore, this gene smoking interaction is biologically plausible. A higher risk of lung cancer among smokers who carried the susceptible genotypes or haplotypes may be attributed to many transformed or preinvasive lung cells caused by tobacco carcinogens, which in turn increases the possibility that one of these cells will become an invasive tumor under the condition of higher lifetime expression of MMP-2. In addition, because MMPs expression can be induced by smoking (39,40), another hypothesis is that in addition to higher constitutive expression of MMPs expression can be induced by smoking (39,40), another hypothesis is that in addition to higher constitutive expression owing to gain of two Sp1 promoter sites, the inducibility of the C−1306C−732 haplotype of MMP2 by smoking may also be higher than that of the T−1306T−732 haplotype, which lacks two Sp1 sites. Given these conditions, it would be expected that subjects who smoked and carried the −1306CC and −735CC genotypes or the C−1306C−732 haplotype were more susceptible to developing lung cancer. Plans have been made to examine the mechanism(s) by which these MMP2 promoter SNPs modify the development of lung cancer.

In summary, this study provides further evidence that the functional genotype and haplotype in the MMP2 promoter are genetic susceptibility factors for lung carcinogenesis. These extended results are consistent with our previous findings in the studies of esophageal cancer, gastric cardia cancer and breast cancer, and further support the hypothesis that gain-of-function of MMP2 resulting from genetic polymorphisms plays an important role in human carcinogenesis.
Acknowledgements

This work was supported by grants 2002BA111A06 from National 863 High Technology Project and grant 2004CB18701 from State Key Basic Research Program.

Conflict of Interest Statement: None declared.

References