Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behavior and tobacco-related lung cancer risk in male Japanese smokers

Masaki Fujieda1,7, Hiroshi Yamazaki1, Tetsuya Saito1, Kazuma Kiyotani1, Maxwell Afari Gyamfi1, Masaharu Sakurai2, Hirotoshi Dosaka-Akita3, Yuichi Sawamura4, Jun Yokota5, Hideo Kunitoh6 and Tetsuya Kamataki1

1Laboratory of Drug Metabolism, Division of Pharmacobiodynamics, Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan, 2Division of Pharmacology and Therapeutics, Graduate School of Clinical Pharmacy, Kumamoto University, Kumamoto 862-0370, Japan, 3Department of Medical Oncology, Graduate School of Medicine, Hokkaido University, Sapporo 060-8638, Japan, 4Maruyama Clinic, Sapporo 060-0812, Japan, 5Department of Medical Oncology, Graduate School of Medicine, School of Clinical Pharmacy, Kumamoto University, Kumamoto 862-0973, Japan, 6Biological Division, National Cancer Center Research Institute, Tokyo 104-0045, Japan and 7Department of Internal Medicine and Thoracic Oncology, National Cancer Center Hospital, Tokyo 104-0045, Japan

We reported previously that subjects homozygous for the cytochrome P450 2A6 (CYP2A6) 4-4 have a lower risk of lung cancer. The purpose of this study was to clarify whether or not the alterations of smoking behavior and risk for lung cancer could be found in subjects possessing novel CYP2A6 variants discovered recently. An epidemiological study was performed with 1094 cases and 611 controls in male Japanese smokers. It was found that the amounts of daily cigarette consumption in subjects who harbored CYP2A6 4-4, 4-10, 7-7, 7-9 and 4-4 genotypes were significantly less than those in subjects carrying the 4-1 genotype (P < 0.01). Even after adjustment with cigarette consumption, the adjusted odds ratios (ORs) for lung cancer were significantly lower in subjects who harbored CYP2A6 4-4, 4-10, 7-7, 7-9, 4-4, 7-7 and 4-4 genotypes than those who possessed the 4-1/1 genotype (P < 0.05). When participants were classified into four groups according to the CYP2A6 genotypes, group 1 (4-1), group 2 (heterozygotes for the 4-1 and a variant allele), group 3 (heterozygotes and homozygotes for variant alleles except for 4-4) and group 4 (4-4), lung cancer risk was found to be less in subjects with the variant of CYP2A6 alleles (group 2, OR of 0.59 [95% confidence interval (CI), 0.44–0.79]; group 3, OR of 0.52 [95% CI, 0.37–0.72]; group 4, OR of 0.30 [95% CI, 0.16–0.57]). The reduced risk for lung cancer was seen more clearly in heavy smokers than in light smokers. Additional stratification analysis showed that the ORs for squamous cell carcinoma (OR of 0.07) and small cell carcinoma (OR of 0.10) were lower than that of adenocarcinoma (OR of 0.39) in group 4. These results suggest that the CYP2A6 is one of the principal determinants affecting not only smoking behavior but also susceptibility to tobacco-related lung cancer.

Introduction

Cytochrome P450 2A6 (CYP2A6) is known as an enzyme responsible for the metabolism of chemicals and drugs such as coumarin (1), nicotine (2), (+)-cis-3,5-dimethyl-2-(3-pyridyl)thiazolidin-4-one hydrochloride (SM-12502) (3), tegafur (4), fadrozole (5), methoxyflurane (6) and valproic acid (7). The enzyme can also metabolically activate a number of carcinogens including tobacco-specific N-nitrosamines, such as (4-methylnitrosamino)-1-(3-pyridyl)-1-butaneone (NNK) (8,9).

The existence of a genetic polymorphism of CYP2A6 was first suggested by evidence that there was a large inter-individual difference in the capacity of coumarin 7-hydroxylation (10,11). In fact, various variants of the CYP2A6 gene have been found in recent years (1,12). Analyzing the genes of subjects who showed a poor metabolizer phenotype toward SM-12502, we found two novel deletion-type variants of the CYP2A6 gene, CYP2A6*4B and CYP2A6*4C (19,20); the CYP2A6*4C was one of the major variants in Japanese. Following the discovery of CYP2A6*4C, we discovered two additional alleles, CYP2A6*7 and CYP2A6*11, showing a decrease in enzymatic activity (21,22). A novel variant, the CYP2A6*9, has a −48T to G nucleotide substitution in the TATA box of the 5′-flanking region of the CYP2A6, which reduced the expression levels of CYP2A6 mRNA and protein in human livers (23). The CYP2A6*10 allele possessing two simultaneous amino acid substitutions seen in the CYP2A6*7 and CYP2A6*8 also shows decreased enzymatic activity (24).

Most cancers are caused by chemical carcinogens present in our environment (25,26). These chemical carcinogens exert their genotoxicity after undergoing metabolic activation by enzymes present in our bodies. Thus, the capacity of enzymes to activate chemical carcinogens has been recognized as one of the factors determining the risk of cancer. Genetic polymorphism of the genes for such enzymes has been expected to be the most typical factor altering the activity and the amounts of the enzymes. Thus, it has been hypothesized that the genetic polymorphism alters the risk of chemical carcinogenesis. However, no conclusive evidence for the association between the genetic polymorphism of carcinogen-activating enzyme and the lung cancer risk has been reported as yet. Recently, several reports, including our group, have demonstrated the role of CYP2A6 genetic polymorphisms in lung cancer risk with some conflicting results in several populations from different ethnicities (27–32). In our previous paper (31), we reported a clear relationship between CYP2A6 genetic polymorphisms and lung cancer risk in smokers. To our knowledge, our results were not supported by other investigators, who reported that no clear association between CYP2A6 genetic polymorphisms and lung cancer risk could be seen (27,29,30,32). The reason for this discrepancy is not known as yet. However, the most possible explanation for this discrepancy is that they analyzed the genes from smokers and non-smokers. Our epidemiology has indicated that the
significant difference can be seen only in smokers, but not in non-smokers. To support our previous results, we needed to clarify further the effects of the CYP2A6 genetic polymorphism on the inter-individual differences in the risk of lung cancer especially in smokers. Thus, we performed a large-scale epidemiological study to investigate the relationship between a variety of genetic polymorphisms of CYP2A6 as well as CYP2A6*4C and tobacco-related lung cancer risk in male Japanese smokers. The present results clearly provide evidence that the variants examined in this study decrease the risk of lung cancer in male Japanese smokers. Based on the results presented in this paper and the results reported quite recently that 8-methoxypsoralen, a specific inhibitor of CYP2A6, completely inhibited the occurrence of adenoma caused by treatment of mice with NNK (33), we propose that genetic polymorphism and the inhibition of CYP2A6 reducing the capacity of CYP2A6 activity result in the reduction of the risk of lung cancer caused by tobacco smoking.

Materials and methods

Subjects

All subjects employed in this study were unrelated male Japanese smokers. Smokers included current and ex-smokers with a smoking history of 0.5 pack/day for at least 1 year. The patient group consisted of a total of 1094 males with a mean age (±SD) of 62.4 ± 9.4 years. The control group consisted of 611 unrelated healthy males with a mean age (±SD) of 53.0 ± 11.2 years. The control subjects did not have any history of cancer. The subjects in case and control groups smoked 53.9 ± 31.9 and 38.3 ± 24.0 pack-years (±SD), respectively. Pack-years smoked was shown to indicate cumulative cigarette dose. Light and heavy smokers were categorized by the 50th percentile pack-years value among controls, and defined as light smokers (<38.3 pack-years) and heavy smokers (≥38.3 pack-years). Most cases were recruited from 1997 to 2003 in the National Cancer Center Hospital, Tokyo, Japan. Incidental cases were ~95% of the total population, with the remaining 5% of prevalent cases. Control subjects recruited in this study composed of healthy volunteers, who visited one of the hospitals that took part in the research within the same time period described above for a health check-up. This procedure provided a natural balance between cases and controls regarding possible confounding factors such as birthplace, since the cancer patients and controls that visited the National Cancer Center Hospital and Maruyama Clinic, respectively, were from all regions of Japan. The age of the lung cancer patients was defined according to the criteria described in the literature pathological classification of lung cancer was determined by more than three pathologists (34). Participants were interviewed using a structured questionnaire, which included questions about the city of residence, birthplace, occupational history, smoking status and previous family (parents and siblings) history of cancer. Each patient in this study was required to sign a consent form. This study was approved by the ethics committee of the National Cancer Center and Hokkaido University.

Genotyping

Genomic DNA was prepared from the peripheral lymphocytes of the patients according to the method of phenol-chloroform extraction followed by ethanol precipitation (35). Genotyping of the CYP2A6*1A, CYP2A6*1B, CYP2A6*4C, CYP2A6*7 and CYP2A6*9 alleles was carried out by the method developed in our laboratory (21,23,36). Genotyping of the CYP2A6*10 allele was carried out by the method reported by Xu et al. (24). The genotyping method, based on the PCR-restriction enzyme fragment length polymorphism (RFLP) for the CYP2A6*4B allele, was newly developed in this study. The reaction mixture (25 μL) for PCR contained LA PCR buffer II, 2.5 mM MgCl2, 2.0 mM dNTPs, 2A6-int4-new sense primer (5'-GCT CCA ATC CAG CCT CGT TTA A-3') and 2A6-in5R-new A5 primer (5'-TCC AGC CCT TGC AGC AAC TG-3') (0.4 μM), 2.0 U of AmpliTag DNA polymerase and ~50 ng of the genomic DNA. PCR was carried out under the following conditions: initial denaturation at 94°C for 1.5 min followed by 30 cycles of reactions composed of cycle denaturation at 94°C for 20 s, annealing at 63°C for 30 s and extension at 72°C for 45 s. After amplification of a 344 bp fragment, restriction digestion using the restriction enzyme, MboII, was carried out. The digested fragments were analyzed on a 1.5% agarose gel. CYP2A6*11/11 yielded 300, 228, 72 and 44 bp fragments, and CYP2A6*11/11 yielded 300 and 44 bp fragments, while CYP2A6*11/1 yielded 228, 72 and 44 bp fragments.

To examine the relationships between the enzymatic function of CYP2A6 predicted by the genotypes and smoking behavior or lung cancer risk, we have defined the following groups: group 1 consists of individuals with two copies of wild-type alleles (CYP2A6*1/1). Group 2 contains subjects carrying at least one wild-type allele (CYP2A6*1/4, *1/7, *1/9, *1/10 or *1/11). Group 3 consists of subjects homozygous or homozygous for variant alleles except for those homozygous for the CYP2A6*4 allele (CYP2A6*4/4, *4/9, *4/10, *4/11, *7/7, *7/9, *7/10, *9/9, *9/10, *9/11 or *10/10). Subjects in the group 4 have two copies of the deletion alleles (CYP2A6*4/4).

Statistical analysis

Differences in age and the amount of cigarette smoking between lung cancer patients and control subjects were tested using unpaired t-test with Welch correction. A relationship between the amount of cigarette smoking and each CYP2A6 genotype was evaluated by one-way ANOVA followed by Games–Howell test as a post hoc comparison. To determine if an association existed between the CYP2A6 genotypes and lung cancer risk, the significance of the difference in the case-control distribution was calculated by χ² test and shown by P value (two-sided). A P value <0.05 was considered to be statistically significant. Compliance with the Hardy–Weinberg equation was tested by χ² test. The association between the genotype distribution and lung cancer risk was assessed by odds ratio (OR) and 95% confidence interval (CI) and were calculated by unconditional logistic regression to adjust for age and cigarette smoking. All statistical computations were carried out using the statistical software SAS, version 5.1 (SAS Institute, Cary, NC).

Results

CYP2A6 genotypes and smoking behavior

Genomic DNA samples from a total of 1705 subjects (1094 lung cancer patients and 611 controls) were analyzed for each CYP2A6 genotype, using the genotyping methods shown in Table I. The relationship between CYP2A6 genotypes and the amounts of daily cigarette consumption in all subjects employed in the present study is shown in Figure 1A. It was found that the amounts of daily cigarette consumption of the subjects, who harbored CYP2A6*4/7, CYP2A6*4/10, CYP2A6*7/7, CYP2A6*7/9 and CYP2A6*4/4, were significantly less than that of the subjects who carried CYP2A6*1/1 (P < 0.01). As mentioned in the Introduction section, the variant CYP2A6 alleles were assumed to generate transcripts possessing lower or no enzymatic activity. Thus, all subjects employed in this study were classified into four groups according to the CYP2A6 genotypes, group 1 (homozygotes for wild-type CYP2A6*1 allele), group 2 (heterozygotes for the CYP2A6*1 allele and a variant allele), group 3 (heterozygotes and homozygotes for variant alleles except for homozygotes for the CYP2A6*4 allele) and group 4 (homozygotes for the CYP2A6*4 allele), to examine the relationship between each genotypic group of CYP2A6 and the amounts of daily cigarette consumption (Figure 1B). Expecting that the CYP2A6 enzymatic function predicted by CYP2A6 genotypes will be related...
to smoking behavior. It was found that the amounts of daily cigarette consumption in the subjects significantly decreased in the order from group 1 to group 4.

**CYP2A6 genotypes and tobacco-related lung cancer risk**

The results of analysis on the association between CYP2A6 genotypes and tobacco-related lung cancer risk are shown in Table II. The CYP2A6*1A and CYP2A6*1B alleles were classified as CYP2A6*1, because the functions of both alleles were thought to be the same. Also, CYP2A6*4B and CYP2A6*4C alleles were classified into CYP2A6*4 (Table II–V), since both alleles caused the same consequence on the enzyme expression. The distribution of CYP2A6 genotypes in both controls and cases were not different from that expected from the Hardy–Weinberg equilibrium ($\chi^2 = 10.1$ and 8.6, $P = 0.90$ and 0.95, respectively) (Table II). In contrast, the distribution of the CYP2A6 genotypes in cases was significantly different from that in controls ($\chi^2 = 42.6, P = 0.0005$). Even after adjustment with cigarette consumption and age by logistic regression analysis, the adjusted ORs for the risk of lung cancer were still significantly lower in subjects who harbored CYP2A6*1/*4, CYP2A6*1/*7, CYP2A6*1/*8, CYP2A6*1/*10, CYP2A6*4/*4, CYP2A6*4/*7, CYP2A6*4/*9, CYP2A6*7/*7 and CYP2A6*7/*9 genotypes as compared with those who possessed the CYP2A6*1/*1 genotype ($P < 0.05$) (Table II).

According to the same criteria as Figure 1B, all subjects employed in the present study were classified into groups 1–4 (Table III). The adjusted ORs of groups 2–4 in overall cases decreased to 0.59 (95% CI of 0.44–0.79), 0.52 (95% CI of 0.37–0.72) and 0.30 (95% CI of 0.16–0.57), respectively. Furthermore, when smokers were stratified by pack-years smoked, the reduced risk of lung cancer was seen more clearly in heavy smokers (smoked ≥38.3 pack-years) according to the genotypes; group 2 (adjusted OR, 0.42; 95% CI, 0.28–0.62), group 3 (adjusted OR, 0.39; 95% CI, 0.25–0.63) and group 4 (adjusted OR, 0.19; 95% CI, 0.05–0.65), than in light smokers (smoked <38.3 pack-years); group 2 (adjusted OR, 0.95; 95% CI, 0.61–1.49), group 3 (adjusted OR, 0.73; 95% CI, 0.46–1.18) and group 4 (adjusted OR, 0.48; 95% CI, 0.22–1.04).

Allele-based analysis on the risk for tobacco-related lung cancer was also performed (Table IV). The distribution of the CYP2A6 alleles between cases and controls was also significantly different ($\chi^2 = 35.7, P < 0.0001$). The ORs were found to be significantly low for the CYP2A6*4, CYP2A6*7, CYP2A6*9 and CYP2A6*10 alleles ($P < 0.05$), supporting the idea that capacity of CYP2A6 activity is one of the determinants affecting susceptibility to tobacco-related lung cancer.

To further examine the impact of the CYP2A6 genetic polymorphism on tobacco-related lung cancer risk, lung cancer patients were divided into three groups, squamous cell carcinoma (SqCC), small cell carcinoma (SCC) and adenocarcinoma (Ad), according to a pathological classification (Table V). SqCC and SCC have been major types of lung cancer caused by smoking, whereas Ad had not been recognized as a common histological type of lung cancer in smokers until recent years, when it was demonstrated that Ad could be increased by smoking (37,38). Significant differences in the distribution of the four CYP2A6 groups between controls and cases suffering from SqCC ($\chi^2 = 20.8, P = 0.0001$), SCC ($\chi^2 = 15.8, P < 0.01$) and Ad ($\chi^2 = 15.6, P < 0.01$) were found. Among overall cases, the adjusted ORs for SqCC (adjusted OR, 0.07; 95% CI, 0.01–0.33) and SCC (adjusted OR, 0.10; 95% CI, 0.01–0.78) were lower than that of Ad (adjusted OR, 0.39; 95% CI, 0.20–0.77) in group 4 (CYP2A6*4/*4). Additional analysis with stratification of histological cancer subtypes revealed that this protection effect for lung cancer was mainly due to markedly reduced risk among heavy smokers. In contrast, among light smokers, no significant association between CYP2A6 groups and the risk for each subtype of lung cancer was found.

**Discussion**

One of the most important issues to clarify was that the lower tobacco-related lung cancer risk seen in smokers possessing the CYP2A6*4C allele appeared in association with the activity of CYP2A6. Thus, we performed an additional epidemiological study to confirm this possibility. In this context, we analyzed the frequency of novel CYP2A6*7, CYP2A6*9, CYP2A6*10 and CYP2A6*11, in addition to CYP2A6*4C, to know if the frequency of these alleles associated with susceptibility to tobacco-related lung cancer. We found that there was a clear relationship between the various CYP2A6 genotypes and tobacco-related lung cancer risk in male Japanese smokers in the present study.

In the Japanese population used in this study, allele frequencies of CYP2A6*1, CYP2A6*4, CYP2A6*7, CYP2A6*9, CYP2A6*10 and CYP2A6*11 in healthy controls were essentially the same as compared with previous studies from our and other laboratories (13,22–24,28,36,39), except for a few reports showing the allele frequencies of CYP2A6*4 in a Chinese (30,32) and CYP2A6*7 and CYP2A6*10 in the Japanese population (40) being 8, 7 and 1%, respectively. These allele frequencies are much lower than that reported by us. The discrepancy of allele frequencies between their
Fig. 1. Relationship between the CYP2A6 genotypes and the amounts of daily cigarette consumption (A), and between CYP2A6 groups classified by CYP2A6 genotypes and daily cigarette consumption (B). Cigarette consumption was investigated with all subjects employed in this study. CYP2A6*1 consists of CYP2A6*1A and *1B alleles. CYP2A6*4 consists of CYP2A6*4B and *4C alleles. The number of subjects is shown under each genotypic group in brackets. Subjects were classified into four groups, according to the CYP2A6 genotypes. Group 1 includes the subjects carrying CYP2A6*1/*1 (wild-type). Group 2 contains subjects heterozygous for wild-type allele (CYP2A6*1/*4, CYP2A6*1/*7, CYP2A6*1/*9, CYP2A6*1/*10, CYP2A6*1/*11). Group 3 consists of subjects carrying CYP2A6*4/*7, CYP2A6*4/*9, CYP2A6*4/*10, CYP2A6*4/*11, CYP2A6*7/*7, CYP2A6*7/*9, CYP2A6*7/*10, CYP2A6*9/*9, CYP2A6*9/*10, CYP2A6*9/*11 or CYP2A6*10/*10. Group 4 contains subjects homozygous for CYP2A6 deletion allele (CYP2A6*4/*4). Genotypes are shown in numerical order. Horizontal lines mean medians. Boxes show 25th and 75th percentile of the observed values. Bars mean 10th and 90th percentiles. The amount of daily cigarette consumption was significantly less than those carrying the CYP2A6*1/*1 genotype (*P < 0.01).

Table II. Distribution of CYP2A6 genotypes in lung cancer patients

<table>
<thead>
<tr>
<th>CYP2A6&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cases (%)</th>
<th>Controls (%)</th>
<th>Crude OR</th>
<th>Adjusted OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 1094)</td>
<td>(n = 611)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1/*1</td>
<td>300 (27.4)</td>
<td>110 (18.0)</td>
<td>1.00&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.00&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>*1/*4</td>
<td>185 (16.9)</td>
<td>94 (15.4)</td>
<td>0.72 (0.52-1.00)</td>
<td>0.68 (0.47-0.98)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*1/*7</td>
<td>106 (9.7)</td>
<td>62 (10.1)</td>
<td>0.63 (0.43-0.92)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.65 (0.46-0.84)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*1/*9</td>
<td>207 (18.9)</td>
<td>112 (18.3)</td>
<td>0.68 (0.49-0.93)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.59 (0.41-0.84)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*1/*10</td>
<td>21 (1.9)</td>
<td>20 (3.3)</td>
<td>0.90 (0.20-0.74)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.30 (0.14-0.61)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*1/*11</td>
<td>9 (0.8)</td>
<td>5 (0.8)</td>
<td>0.66 (0.22-2.01)</td>
<td>1.05 (0.29-3.78)</td>
</tr>
<tr>
<td>*4/*4</td>
<td>25 (2.3)</td>
<td>28 (4.6)</td>
<td>0.33 (0.18-0.59)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.29 (0.15-0.56)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*4/*7</td>
<td>36 (3.3)</td>
<td>30 (4.9)</td>
<td>0.44 (0.26-0.75)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.45 (0.25-0.82)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*4/*9</td>
<td>62 (5.7)</td>
<td>44 (7.2)</td>
<td>0.52 (0.33-0.81)&lt;sup&gt;b, d&lt;/sup&gt;</td>
<td>0.51 (0.31-0.83)&lt;sup&gt;b, d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*4/*10</td>
<td>16 (1.5)</td>
<td>16 (2.6)</td>
<td>0.37 (0.18-0.76)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.55 (0.24-1.28)</td>
</tr>
<tr>
<td>*4/*11</td>
<td>2 (0.2)</td>
<td>2 (0.3)</td>
<td>0.37 (0.05-2.64)</td>
<td>0.52 (0.05-5.62)</td>
</tr>
<tr>
<td>*7/*7</td>
<td>12 (1.1)</td>
<td>13 (2.1)</td>
<td>0.34 (0.15-0.76)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.28 (0.11-0.71)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*7/*9</td>
<td>39 (3.6)</td>
<td>32 (5.2)</td>
<td>0.45 (0.27-0.75)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.44 (0.25-0.79)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>*7/*10</td>
<td>7 (0.6)</td>
<td>4 (0.7)</td>
<td>0.64 (0.18-2.24)</td>
<td>0.98 (0.25-3.92)</td>
</tr>
<tr>
<td>*9/*9</td>
<td>55 (5.0)</td>
<td>30 (4.9)</td>
<td>0.67 (0.41-1.10)&lt;sup&gt;b, d&lt;/sup&gt;</td>
<td>0.71 (0.41-1.23)</td>
</tr>
<tr>
<td>*9/*10</td>
<td>11 (1.0)</td>
<td>4 (0.7)</td>
<td>1.01 (0.31-3.23)</td>
<td>0.98 (0.28-3.41)</td>
</tr>
<tr>
<td>*9/*11</td>
<td>1 (0.1)</td>
<td>1 (0.2)</td>
<td>0.37 (0.02-5.91)</td>
<td>0.36 (0.02-5.81)</td>
</tr>
<tr>
<td>*10/*10</td>
<td>0 (0.0)</td>
<td>4 (0.7)</td>
<td>NA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>NA&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Significant difference in the distribution of CYP2A6 genotypes was found between lung cancer cases and control subjects (χ² value 42.6, P = 0.0005). *CYP2A6*1 consists of CYP2A6*1A and *1B alleles. CYP2A6*4 consists of CYP2A6*4B and *4C alleles.

<sup>a</sup>To adjust age and smoking habit, OR and 95% CI were calculated by logistic regression.

<sup>b</sup>Reference category.

<sup>c</sup>Significant decrease of OR is indicated by 95% CI.

<sup>d</sup>Not applicable.

The ethnic difference of the CYP2A6*4 allele frequency might exist between Japanese and Chinese populations. In fact, the frequencies of CYP2A6*4 in controls in their two different studies were almost the same (30,32), although the frequencies in lung cancer cases were largely different.

We assessed the impact of CYP2A6 genetic polymorphisms on the number of cigarettes smoked per day and the risk of lung cancer. Regarding the basis for the classification of the CYP2A6 genotypes into groups 1-4, we recently analyzed that the relationship between the in vivo catalytic activity of CYP2A6 towards nicotine and the polymorphism of the CYP2A6 gene in healthy Thai volunteers (unpublished data). The levels of plasma cotinine concentration in subjects genotyped as CYP2A6*1/*4, CYP2A6*1/*7, CYP2A6*1/*9, CYP2A6*1/*10, CYP2A6*4/*7, CYP2A6*4/*9, CYP2A6*7/*7, and CYP2A6*9/*9 showed 53.9, 61.4, 72.2, 63.4, 11.7, 35.8, 20.4, and 58.9% of the plasma cotinine concentration of subjects carrying CYP2A6*1/*1, respectively, suggesting that the catalytic activity of CYP2A6 is lower in the subjects homozygous for either CYP2A6*7 or CYP2A6*9, or heterozygous within the CYP2A6*4/*4, CYP2A6*7, CYP2A6*9/*9, and CYP2A6*10 variants. Additionally, we also found the CYP2A6*4/*11 genotype from a patient who also showed a poor metabolic phenotype in the metabolism of tegafur to yield 5-fluorouracil (22), suggesting that the enzyme encoded by CYP2A6*11 had a lower metabolic capacity. In fact, we clarified that the recombinant CYP2A6.11 had a lower capacity to metabolize tegafur (41% of CYP2A6.1) and coumarin (59%) (22). Furthermore, analyzing the plasma concentration of nicotine, Xu and colleagues (24) have reported that individuals who possessed the CYP2A6*7/*7, CYP2A6*4/*7 and CYP2A6*4/*10 genotypes showed apparently intermediate and poor metabolic phenotype, probably indicating that the CYP2A6*7 and CYP2A6*10 are among the causative alleles.
reducing enzymatic activity of CYP2A6. These lines of evidence support our idea that not only the amounts of daily cigarette consumption but also the risk for tobacco-related lung cancer decrease in association with the impaired function of CYP2A6 (Figure 1B and Table III). Results reported by Tyndale and coworkers (41–44) on the association between the CYP2A6 genetic polymorphism with smoking behavior are in agreement with our results. The present study also clearly indicates that the predicted capacity of CYP2A6 correlates well with the tobacco-related lung cancer risk, suggesting that the inhibition of this enzyme by some inhibitors of this enzyme results in the prevention of the occurrence of tobacco-related lung cancer. Supporting this idea, our recent study showed that treatment of A/J mice with NNK together with 8-methoxypсорalen, a specific inhibitor of CYP2As, completely abolished the occurrence of NNK-induced adenoma (33).

SqCC and SCC have been recognized as major types of lung cancer caused by smoking, whereas Ad has not been regarded as a common histological type of lung cancer caused by smoking until recent years, when it was demonstrated that Ad could be increased by smoking (37,38). Thus, it is of interest to note that in the present study the decreased ORs are seen in SqCC and SCC rather than in Ad, which was in agreement with a previous concept that SqCC and SCC appeared highly related to tobacco smoking.

Conflicting results have been reported on the association of CYP2A6 genetic polymorphisms and lung cancer risk (27–32). These contradictory results seem to be caused by several factors. First, the original genotyping method (45), which was employed in the previous two reports (27,46) is rather non-specific, which caused a misclassification of CYP2A6 genotypes. Secondly, the frequencies of the inactive alleles such as CYP2A6*2 and CYP2A6*4 in their studies were too small to detect a potential relationship with sufficient statistical power (29,30). A larger population is needed to confirm their findings. Thirdly, they analyzed the genotypes of combined groups of smokers and non-smokers (30,32) as pointed out in the Introduction section. As reported in this and a previous paper (31), we found that the association between the genotype of CYP2A6 and the lung cancer risk could be seen only in smokers. In our preliminary results, ORs of subjects heterozygous for the wild-type allele (CYP2A6*1/*4, CYP2A6*1/*7, CYP2A6*1/*9, CYP2A6*1/*10, CYP2A6*1/*11) Group 3 consists of subjects carrying CYP2A6*4/*7, CYP2A6*4/*9, CYP2A6*4/*10, CYP2A6*4/*11, CYP2A6*7/*7, CYP2A6*7/*9, CYP2A6*7/*10, CYP2A6*9/*9, CYP2A6*9/*10, CYP2A6*9/*11 or CYP2A6*10/*10. Group 4 contains subjects homozygous for CYP2A6 deletion allele (CYP2A6*4/*4).

Table IV. Allele frequency of CYP2A6 in lung cancer patients

<table>
<thead>
<tr>
<th>Allele</th>
<th>Cases (%)</th>
<th>Controls (%)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2A6*1</td>
<td>1128 (51.6)</td>
<td>513 (42.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>CYP2A6*4</td>
<td>351 (16.0)</td>
<td>242 (19.8)</td>
<td>0.66 (0.54-0.80)</td>
</tr>
<tr>
<td>CYP2A6*7</td>
<td>212 (9.7)</td>
<td>154 (12.6)</td>
<td>0.63 (0.50-0.79)</td>
</tr>
<tr>
<td>CYP2A6*9</td>
<td>430 (19.7)</td>
<td>253 (20.7)</td>
<td>0.77 (0.64-0.93)</td>
</tr>
<tr>
<td>CYP2A6*10</td>
<td>55 (2.5)</td>
<td>52 (4.3)</td>
<td>0.48 (0.32-0.71)</td>
</tr>
<tr>
<td>CYP2A6*11</td>
<td>12 (0.5)</td>
<td>8 (0.7)</td>
<td>0.68 (0.28-1.68)</td>
</tr>
</tbody>
</table>

Significant difference in the distribution of the six alleles between lung cancer cases and control subjects was found (x2 value 35.7, P < 0.0001).

Table III. Relationship between the CYP2A6 groups and lung cancer risk

<table>
<thead>
<tr>
<th>Group</th>
<th>Cases (%)</th>
<th>Controls (%)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>OR (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>300 (27.4)/110 (18.0)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>528 (48.3)/293 (47.9)</td>
<td>0.59 (0.44-0.79)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>241 (22.0)/180 (29.5)</td>
<td>0.52 (0.37-0.72)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25 (2.3)/28 (4.6)</td>
<td>0.30 (0.16-0.57)</td>
<td></td>
</tr>
</tbody>
</table>

These contradictory results seem to be caused by several factors. First, the original genotyping method (45), which was employed in the previous two reports (27,46) is rather non-specific, which caused a misclassification of CYP2A6 genotypes. Secondly, the frequencies of the inactive alleles such as CYP2A6*2 and CYP2A6*4 in their studies were too small to detect a potential relationship with sufficient statistical power (29,30). A larger population is needed to confirm their findings. Thirdly, they analyzed the genotypes of combined groups of smokers and non-smokers (30,32) as pointed out in the Introduction section. As reported in this and a previous paper (31), we found that the association between the genotype of CYP2A6 and the lung cancer risk could be seen only in smokers. In our preliminary results, ORs of subjects heterozygous for the wild-type allele (CYP2A6*1/*4, CYP2A6*1/*7, CYP2A6*1/*9, CYP2A6*1/*10, CYP2A6*1/*11) Group 3 consists of subjects carrying CYP2A6*4/*7, CYP2A6*4/*9, CYP2A6*4/*10, CYP2A6*4/*11, CYP2A6*7/*7, CYP2A6*7/*9, CYP2A6*7/*10, CYP2A6*9/*9, CYP2A6*9/*10, CYP2A6*9/*11 or CYP2A6*10/*10. Group 4 contains subjects homozygous for CYP2A6 deletion allele (CYP2A6*4/*4).
Tobacco smoke contains a number of tobacco-specific N-nitrosamines, such as N-nitrosodiethylamine, NNK and N'-nitrosornornicotine (48). In addition to CYP2A6, CYP1A1 and CYP2A13 are able to activate NNK (8,9,49,50). Thus, it can be expected that the genetic polymorphism of the CYP1A1 and CYP2A13 genes affect the tobacco-related lung cancer risk. Recently, we found 14 novel CYP2A13 haplotypes including the Arg257Cys variant, which was named as CYP2A13*2 (51). Wang et al. (32) have reported recently that the frequency of the CYP2A13 variant associated with the reduced risk of lung Ad in light smokers. However, analyzing our data using the same subjects employed in the present study, we found no clear association between the lung cancer risk and the CYP2A13*2 allele (data not shown). The reason for this discrepancy is unknown at present. Furthermore, the contribution to cancer risk of other carcinogens such as polycyclic aromatic hydrocarbons and aromatic amines in tobacco smoke could not be ruled out. In fact, the enzymes belonging to the CYP1 gene family play central roles in the metabolic activation of these compounds present in tobacco smoke (52,53). However, we were unable to find out any clear relationships between genetic polymorphism of CYP1A1 and tobacco-related lung cancer risk with the same population employed in the previous epidemiological study (31), probably suggesting that the metabolic activation by CYP2A6 of nitrosamines or carcinogens other than polycyclic aromatic hydrocarbons is the key step determining the tobacco-related lung cancer risk.

In conclusion, our results suggest strongly that the genetic polymorphism of CYP2A6 is one of the principal determinants affecting not only smoking behavior but also tobacco-related lung cancer risk in the Japanese population.

Acknowledgements

This work was supported in part by a Grant-in-Aid (No. 99-2) from the Organization for Pharmaceutical Safety and Research (OPSR) and Ministry of Education, Science, Sports and Culture of Japan. This work was also supported in part by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan, by a Grant-in-Aid from the Core Research for Evolutional Science and Technology, and an SRF Grant for Biomedical Research in Japan.

References

Genetic polymorphism of CYP2A6 and lung cancer risk


Received May 12, 2004; revised June 29, 2004; accepted August 2, 2004