Role of Polymorphisms in Codons 143 and 160 of the $O^6$-Alkylguanine DNA Alkyltransferase Gene in Lung Cancer Risk

Tajinder B. Kaur, John M. Travaline, John P. Gaughan, John P. Richie, Jr., Steven D. Stellman, and Philip Lazarus

Departments of Pathology and Laboratory Medicine [T. B. K., P. L.], Pulmonary Medicine [J. M. T.], and Biostatistics [J. P. G.], Temple University School of Medicine, Philadelphia, Pennsylvania 19140, and Divisions of Nutritional Carcinogenesis [J. P. R.] and Epidemiology [S. S.], American Health Foundation, Valhalla, New York 10595

Abstract

$O^6$-Alkylguanine DNA alkyltransferase (AGT) plays an important role in the repair of alkylating agent-induced DNA damage and protection from the carcinogenic effects of environmental agents. To examine the importance of the AGT codon 143 and codon 160 polymorphisms in risk for lung cancer and to assess the prevalence of these polymorphisms in different racial groups, we performed genotype analysis of lung cancer patients and matched controls. The prevalence of the AGT$^{143Val}$ allele in controls was 0.07 in Caucasians and 0.03 in African Americans. The AGT$^{143Val}$ allele was not detected in an unmatched Asian control cohort. The prevalence of the AGT$^{160Arg}$ variant allele was 0.02 in African Americans, and 0.03 in Asians. A marginally significant association was observed between the AGT codon 143 (isoleucine/valine) genotype and risk for lung cancer (odds ratio = 2.1; 95% confidence interval = 1.01–4.7). The prevalence of the AGT$^{160Arg}$ variant allele was similar in lung cancer cases versus matched controls. These results suggest that the AGT codon 143 polymorphism may play an important role in risk for lung cancer.

Introduction

The AGT$^3$ DNA repair enzyme repairs $O^6$-alkylguanine adducts induced in DNA by a variety of alkylating agents, including environmentally important $N$-nitroso compounds such as dimethylnitrosamine and $N$-methyl-$N'$-nitro-$N'$-nitrosoguanidine (1–3). Cells lacking AGT are more sensitive to mutagenesis and demonstrate increased cytotoxicity when exposed to alkylating agents (2, 4, 5). The repair of alkylguanine adducts may be directly correlated with cellular expression or activity of AGT (2, 6–9). Increased AGT activity has been associated with decreased susceptibility to tumor induction by alkylating agents in transgenic mice (10, 11). DNA adducts induced by the tobacco-specific nitrosamine NNK are substrates for AGT (12, 13), and the defective repair of $O^6$-alkylguanine by AGT in target cells and tissues may play a role in tumor induction in rodents treated with NNK (14). These data are consistent with the fact that reduced capacity to repair $O^6$-methylguanine adducts has been correlated with increased risk for lung cancer (7).

The human AGT is a ubiquitous protein that irreversibly transfers an alkyl group from the $O^6$ position of an adducted guanine to an internal cysteine residue at position 145 of the AGT protein. Mutations at the cysteine acceptor site or at positions in close proximity to this site can inactivate or significantly decrease AGT activity (3, 15–17). Two AA-altering polymorphisms in close proximity to the cysteine acceptor site (at codons 143 and 160) have recently been identified in the AGT gene (18, 19). Although functional studies have not yet been performed for the AGT$^{143Val}$ allele, this polymorphism is located almost adjacent to the alkyl acceptor Cys$^{145}$ in the active site of the AGT enzyme. The polymorphism in codon 160 results in a Gly$\rightarrow$Arg AA change in the AGT protein that codes for an AGT$^{160Arg}$ variant isoform, which exhibits significantly decreased activity in the repair of bulky alkylated DNA adducts as compared with the AGT$^{160Gly}$ wild-type isoform (20). The overall goal of the present study was to assess the polymorphic prevalence of the codon 143 (Ile/Val) and codon 160 (Gly/Arg) genotypes in different racial groups and to examine the potential importance of these polymorphisms as potential risk factors for lung cancer.

Materials and Methods

Subjects. This study was a retrospective case-control study using a 1:1 matched pair design for the purposes of minimizing differences due to age, sex, and race. Based on preliminary studies suggesting that the prevalence of the AGT codon 143 (Ile/Val) genotype is 21% (19), power calculations using a 1:1 matching showed that 94 cases would be required to detect an OR of 2.5 for the AGT codon 143 (Ile/Val) genotype with 80% power at the 0.05% significance level. Due to the variation in the prevalence of the AGT codon 160 (Gly/Arg) genotype in different studies (18, 19, 21), no power calculations were made for this polymorphism.

Cases ($n = 139$) were Caucasian or African American patients diagnosed with primary lung cancer and were identified between 1997 and 1999 at TUH (Philadelphia, PA). Diagnosis of lung cancer was provided by pathology confirmation (performed in the TUH Pathology Department), and all pathology-confirmed lung cancer patients who visited the Pulmonary Division at TUH were considered potentially eligible cases for our study. All of the cases entered into the study were recruited.
within 11 months after diagnosis, with 93% of all eligible cases recruited into the study. Of those eligible cases not recruited, three died before potential contact due to advanced disease. The histological distribution of lung cancer types was as follows: (a) adenocarcinoma (n = 38); (b) squamous cell carcinoma (n = 37); (c) small cell carcinoma (n = 31); (d) non-small cell carcinoma (n = 25); and (e) others (n = 8). Controls were individually matched to cases and were comprised of inpatients treated at TUH whose admitting diagnosis was unrelated to smoking. Exclusions for controls included cardiovascular disease, peripheral vascular disease, stroke and syncope, cirrhosis, emphysema, chronic bronchitis, chronic obstructive pulmonary disease, and other tobacco-related cancers (including bladder, esophageal, laryngeal, oral cavity, pancreatic, liver, and kidney cancer). The main diseases diagnosed in the matched control subjects were trauma and other ambulatory dysfunctions (34%), musculoskeletal (19%), gastrointestinal (11%), respiratory (10%), infectious (8%), and renal (6%). Eligible controls were, on average, recruited within 2 months of case recruitment by screening of TUH admission rosters for inpatients. The eligible control pool was restricted to subjects with the same age at diagnosis (≥ 5 years), race, and sex as cases. A detailed questionnaire containing questions on demographics and lifelong smoking habits was administered to all subjects. Tobacco use was categorized as described previously (22). Smokers were classified as having smoked at least one cigarette/day for a minimum of 1 year (0.05 py). As a separate control group, we recruited 36 noncancerous, “healthy” TUH employees of Asian descent. These subjects were recruited to assess the prevalence of AGT polymorphic alleles in the Asian population and were not included in the lung cancer case-control study described above.

**Genotype Analysis.** Buccal cell samples were collected from all subjects for the analysis of polymorphic genotypes, as approved by the TUH institutional review board. Informed consent was obtained from all subjects, with a consent rate of 98% for cases and 80% for controls who were asked to participate. Buccal cell samples were collected and processed, and DNA was isolated as described previously (22). Because AGT codons 143 and 160 are both within exon 5 of the AGT gene, a single PCR amplification was performed for the screening of both codon 143 and codon 160 polymorphisms. PCRs were performed using 300 ng of sense (5′-TGACCCCTTCAGTTCCCAT-3′) and antisense (5′-CAGTGTGTGTCGCTTAAAC-3′) primers homologous to exon 5 sequences of the AGT gene to generate a 262-bp fragment. A standard PCR was performed as described previously (22) with a primer annealing step of 58°C for 45 s. A separate PCR amplification was also performed for a DNA specimen (chosen at random) using a codon 143 Val-encoding sense primer (5′-TGACCCCTTCAGTTCCCAT-3′) that contains a guanine (underlined nucleotide in the primer sequence) at the polymorphic nucleotide site of AGT codon 143. This PCR amplimer was sequenced to confirm AGT genotyping results. DNA sequencing to confirm AGT genotyping results. Prevalence of AGT 143Val and AGT 160Arg alleles in controls

<table>
<thead>
<tr>
<th></th>
<th>AGT160Arg</th>
<th>AGT143Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasians</td>
<td>53</td>
<td>0.05</td>
</tr>
<tr>
<td>African Americans</td>
<td>80</td>
<td>0.19</td>
</tr>
<tr>
<td>Asians</td>
<td>35</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 1 Prevalence of AGT143Val and AGT160Arg alleles in controls

New England Nuclear, Boston, MA) for each low-melting point agarose-purified PCR-amplified DNA sample. Samples were subjected to one incubation cycle consisting of 95°C for 3 min, 60°C for 30 s, and 72°C for 40 s before sample processing and electrophoresis. All genotyping analysis was repeated for at least 20% of the specimens, and selected PCR-amplified DNA samples (n = 2 for both the wild-type and heterozygous genotypes for each polymorphism) were examined by dideoxy DNA sequencing to confirm AGT genotyping results.

**Statistics.** The risk of lung cancer in relation to AGT genotype and smoking was estimated using conditional logistic regression appropriate for a 1:1 matched pair design to calculate ORs and 95% CIs, adjusting for age, sex, race, and py. The potential association between AGT polymorphisms and lung cancer risk was analyzed for informative case-control pairs using the Pearson’s χ² test analysis. To analyze the potential association between AGT genotype and lung cancer risk after stratification into smoking categories, multiple logistic regression analysis was performed after adjusting for age, sex, race, and py. Student’s t test (two-tailed) was used for comparing continuous variables such as smoking between cases and controls. All Ps are presented as two-sided. The statistical computer software SAS/STAT (version 6.12, fourth edition, Vol. 2; SAS Institute, Cary, NC) was used to perform all statistical analysis.

**Results and Discussion**

A total of 56 Caucasian and 83 African-American lung cancer cases and an equal number of individually matched controls were entered into the study. The average age was identical for cases and controls (63 years for African Americans and 66 years for Caucasians), and the male:female ratio was virtually the same for African Americans (55% male) and Caucasians (57% male). For Caucasians, 89% of cases and 71% of controls were ever-smokers (i.e., <0.05 py), whereas 99% of cases and 85% of controls were ever-smokers in the African-American group. Cigarette consumption was significantly higher in cases (48 py for African Americans and 52 py for Caucasians) as compared to controls (30 py for African Americans and 32 py for Caucasians) for both racial groups (P < 0.005 for Caucasians and P < 0.002 for African Americans). The mean age of the unmatched, noncancerous Asian cohort was 37 years, and 58% of these subjects were male. This Asian cohort (n = 36) was 69% Chinese, 14% Korean, 11% Vietnamese, 3% Filipino, and 3% Japanese.

The AGT160Arg allelic prevalence in controls was 0.009 for Caucasians and 0.019 for African Americans (Table 1). The AGT160Arg allele was observed in 6% of the Asian cohort (allelic prevalence = 0.029). Both of the Asian subjects with a detectable AGT160Arg allele were of Chinese descent. The prevalence of the AGT143Val allele in controls was 0.073 for Cau-
A marginally significant association was observed between the AGT codon 143 (Ile/Val) genotype and increased risk for lung cancer, regardless of whether this analysis was performed by analyzing crude genotyping data (\( \chi^2 = 3.98; P = 0.047 \)) or by conditional logistic regression analysis (OR = 2.1, 95% CI = 1.01–4.7; Table 2). Similar increases in risk were observed for African Americans (OR = 2.3, 95% CI = 0.73–8.3) and Caucasians (OR = 2.0, 95% CI = 0.78–5.7). A similar prevalence of the AGT143Val allele was observed for both lung cancer cases and controls for African Americans and Caucasians.

To examine the effects of smoking, the AGT143Val allele and lung cancer risk were examined separately in smokers. Smokers [never-smokers (n = 7) were excluded from this analysis] were categorized into two groups based on lifetime smoking history divided at the median py for all subjects (32 py). Similar risk was observed in both the light-smoking (≤32 py; OR = 1.6, 95% CI = 0.33–7.3) and heavy-smoking (>32 py; OR = 1.7, 95% CI = 0.61–4.7) groups. A test for interaction between genotype and smoking was not significant in our conditional logistic regression model for lung cancer risk.

Many studies have examined the role of genetic polymorphisms in susceptibility to environmental carcinogens. The majority of these studies have focused on polymorphisms present in genes coding for enzymes involved in carcinogen metabolism. Few studies have reported on DNA repair as a mechanism for differential cancer susceptibility. Codon 143 in the AGT gene encodes an AA located within the active site of the AGT protein, two AAs upstream of the alkyl acceptor Cys145. The Ile143 is conserved among all known mammalian AGTs, and although an Ile→Val AA change similar to that observed at AGT codon 143 is conservative in nature, previous studies have shown that the polymorphic Ile→Val AA change at codon 105 of the glutathione S-transferase P1 gene is associated with decreased catalytic activity of the variant glutathione S-transferase π protein (24). In addition, although not in the active site of the AGT protein, the genetically linked codon 178 Lys→Arg AA change within the AGT143Val variant may alter the function of this protein. Function-activity relationships of the AGT143Val variant must be established for both the codon 143 and codon 178 polymorphic sites to better assess their potential importance in relation to lung cancer risk.

The association between the AGT143Val variant and lung cancer risk observed in this study suggests that decreased activity in the repair of adducts generated by tobacco smoke carcinogens confers increased risk for lung cancer for subjects with this variant allele. Although no difference in attributable risk was observed in subjects at different smoking doses, the AGT143Val variant may interact with tobacco smoke carcinogens to confer risk in all smokers. It is not known whether the AGT143Val variant exhibits altered activity toward the repair of specific types of adducts similar to that observed for the AGT160Arg variant (20). If the AGT143Val variant exhibits altered repair activity against methylated adducts, such studies could potentially link AGT polymorphisms and lung cancer susceptibility with adducts induced by alkylating tobacco carcinogens such as NNK (14). Because only seven cases in our study were never-smokers, a larger study will be required to examine potential associations in smokers versus never-smokers.

In a recent study, the AGT143Val allele was detected in 21% of a small cohort of noncancerous Caucasian controls but was not detected in 60 noncancerous Chinese controls (19). A difference in AGT143Val allelic prevalences between racial groups was also observed in the present study. The difference in AGT143 genotype prevalence between these two studies is likely due to the low number of Caucasian subjects (n = 28) screened in the earlier study (19). By contrast, variability in the prevalence of the AGT160Arg allele has been observed between studies of different cohorts of the same race. Imai et al. (18) described the presence of this allele in 3 of 28 healthy Japanese controls. However, this allele was not detected in two studies examining allelic prevalence in Caucasians (19, 21), African Americans (21), and Asians [including Chinese (19) and a small cohort of Japanese (21)]. In the present study, we confirm the existence of the AGT160Arg allele in humans; however, it appears to be present at a low prevalence (i.e., ≤0.03) in multiple racial groups. This is consistent with studies demon-
strating the lack of a 6-benzylguanine-resistant AGT phenotype in Caucasians and Japanese (25). Although the data from the present study suggest that the AGT<sup>160Arg</sup> allele is not associated with a high degree of attributable risk for lung cancer, the low prevalence of this allele necessitates that a much larger study be performed to obtain sufficient power for an efficient risk assessment analysis.

**Acknowledgments**

We thank Jong Park for critical comments and review of the manuscript, Mark Pietras and Megan Grossman for excellent technical assistance, and Donald Pollard and Michelle Hayes for assistance in subject identification and recruitment.

**References**


11. Nakatsuka, Y., Matsukuma, S., Nemoto, N., Sagano, H., Sekiguchi, M., and Ishikawa, T. 6-Methylguanine DNA methyltransferase protects against nitro-