Tobacco Carcinogen-Detoxifying Enzyme UGT1A7 and Its Association With Orolaryngeal Cancer Risk

Zhong Zheng, Jong Y. Park, Chantal Guillemette, Stimson P. Schantz, Philip Lazarus

Background: UDP-glucuronosyltransferase 1A7 (UGT1A7) detoxifies several tobacco carcinogens. We determined whether UGT1A7 expression is observed in normal orolaryngeal tissue and whether UGT1A7 allelic variations are associated with the risk for orolaryngeal cancer. Methods: UGT1A7 expression in normal orolaryngeal tissue was determined by semiquantitative reverse transcription–polymerase chain reaction (PCR). Buccal cell DNA isolated from 194 case subjects with orolaryngeal cancer and from 388 control subjects who were matched by sex, age, and race was subjected to UGT1A7 genotyping with the use of combined PCR–restriction fragment length polymorphism and allelic discrimination analysis. All statistical tests were two-sided. Results: UGT1A7 messenger RNA was expressed at similar levels in the esophagus, tongue, tonsil, floor of the mouth, and larynx. Genotyping revealed the presence of three variant reduced-activity UGT1A7 alleles in both Caucasians and African-Americans. Individuals with any of the predicted low-activity UGT1A7 genotypes had an increased risk of orolaryngeal cancer (odds ratio [OR] = 3.7; 95% confidence interval [CI] = 1.7 to 8.7) relative to subjects with the wild-type genotype. Both Caucasians and African-Americans with the low-activity genotypes had statistically significantly increased orolaryngeal cancer risk compared with Caucasians and African-Americans with the wild-type genotype (OR = 2.8 [95% CI = 1.1 to 7.6] and OR = 6.2 [95% CI = 1.2 to 31], respectively). For subjects with the predicted low-activity genotypes, the risks of oral cavity cancer (OR = 4.2; 95% CI = 1.7 to 10) and laryngeal cancer (OR = 3.7; 95% CI = 0.99 to 14) were similar. There was no association between UGT1A7 genotype and orolaryngeal cancer risk in never smokers, whereas subjects with predicted low-activity UGT1A7 genotypes who were light smokers (OR = 3.7; 95% CI = 1.1 to 12) or heavy smokers (OR = 6.1; 95% CI = 1.5 to 25) had an increased risk. Conclusions: The tissue expression of UGT1A7 is consistent with the possibility of a physiologic role in orolaryngeal cancer. Variations in the UGT1A7 gene that reduce UGT1A7 activity may affect the risk of smoking-related orolaryngeal cancer. [J Natl Cancer Inst 2001;93:1411–8]

UDP-glucuronosyltransferases (UGTs) are a family of enzymes that metabolize endogenous compounds, such as bilirubin and steroid hormones, drugs, and environmental carcinogens by glucuronidation (1). UGT2B7, UGT1A9, and UGT1A7 are UGTs that have been implicated in the conjugation and detoxification of the tobacco carcinogens 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (2,3) and benzo[a]pyrene (BaP) (4,5). Although UGT1A7, UGT1A9, and UGT2B7 are all expressed in the esophagus (6), it is not known whether these UGTs are expressed in other aerodigestive tract tissues.

Many genes that encode enzymes involved in the metabolic activation or detoxification of carcinogens contain polymorphisms. A growing body of evidence suggests that many of these genetic polymorphisms are associated with the risk of developing cancers of the aerodigestive tract (7,8). Although many molecular epidemiologic studies have been performed to assess risks associated with the polymorphisms found in many of these enzymes, few studies have been performed for the UGTs. One exception is the “TATA” box polymorphism in the promoter region of UGT1A1 commonly associated with Gilbert’s syndrome (9), which has recently been implicated in the risk for breast cancer (10). Although polymorphisms have been identified in genes encoding UGT2B4, UGT2B7, UGT2B15, and UGT1A6 (11–14), few studies have examined their effects on enzyme function, and no studies have examined their importance in cancer risk.

In studies examining polymorphisms in the UGT1A7 gene, three allelic variants of the wild-type UGT1A7*1 allele have been described (5). Two of the polymorphisms, present in the UGT1A7*2 and UGT1A7*3 alleles, are genetically linked and result in asparagine-to-lysine and arginine-to-lysine amino acid substitutions at codons 129 and 131, respectively. An additional polymorphism at codon 208, which is present in the UGT1A7*3 and UGT1A7*4 alleles, is unlinked to the polymorphisms at codons 129 and 131 and results in the substitution of arginine for tryptophan. The UGT1A7 isozyme encoded by the wild-type allele (UGT1A7*1) is twofold to fourfold more active against the BaP metabolites 7-, 8-, and 9-hydroxy-BaP than the enzymes encoded by the three variant alleles UGT1A7*2, UGT1A7*3, and UGT1A7*4 (5).

Because UGT1A7 plays an important role in the detoxification of several tobacco carcinogens, we examined the potential role of UGT1A7 genotype in the risk for orolaryngeal cancer by assessing the expression pattern of UGT1A7 in different orolaryngeal tissues and the genotypes of individuals with and without orolaryngeal cancer.

MATERIALS AND METHODS

Tissue Samples and Analysis of UGT1A7 Expression

Samples (2–6 g) of normal human liver, esophagus, larynx, tongue, tonsil, and floor of mouth were obtained from the Tissue Procurement Facility at the H. Lee Moffitt Cancer Center and Research Institute (Tampa, FL). All of the specimens came from subjects undergoing cancer surgery. Liver specimens were obtained from subjects who had never smoked, whereas all of the orolaryngeal and esophageal specimens were obtained from recent smokers. All of the tissue samples were quick-frozen at −70°C within 30 minutes of surgery. Total RNA was isolated from 0.5-g samples of normal liver and orolaryngeal tissues with the use of the guanidinium isothiocyanate/cesium chloride method, followed by treatment with deoxyribonuclease I as described previously (15). Total RNA was stored in aliquots at −70°C.

Affiliations of authors: Z. Zheng, J. Y. Park, P. Lazarus, Divisions of Cancer Control and Molecular Oncology, H. Lee Moffitt Cancer Center and Research Institute, Interdisciplinary Oncology Program and Departments of Biochemistry and Pharmacology and Therapeutics, University of South Florida, Tampa; C. Guillemette, Oncology and Molecular Endocrinology Research Center, CHUL Research Center (CHUQ), School of Pharmacy, Laval University, PQ, Canada; S. P. Schantz, Department of Otolaryngology, The New York Eye and Ear Infirmary, New York City.

Correspondence to: Philip Lazarus, Ph.D., Divisions of Cancer Control and Molecular Oncology, H. Lee Moffitt Cancer Center and Research Institute, University of South Florida, MRC-2E, 12902 Magnolia Dr., Tampa, FL 33612 (e-mail: plazarus@moffitt.usf.edu).

See “Notes” following “References.”

© Oxford University Press
Three micrograms of total RNA was reversed transcribed in a volume of 20 μL that contained 200 U of Moloney murine leukemia virus reverse transcriptase (Life Technologies Inc., [GIBCO/BRL], Rockville, MD) and 0.5 μg of oligo (dT)12-18 primer, as outlined in the manufacturer’s protocol. Polymerase chain reaction (PCR) was performed in a volume of 50 μL that contained 5 μL of reverse transcription (RT) reaction mix, 2.4 mM MgCl2, 50 mM KCl, 20 mM Tris–HCl (pH 8.0), 0.2 mM of each deoxynucleoside triphosphate, 400 pmol each of the sense and antisense UGT primers, and 2.5 U of Taq DNA polymerase (Boehringer Mannheim, Indianapolis, IN). Sense (1A7s1: 5′-ATG-CTCGCTTGAGCGACCACTTG-3′; corresponding to nucleotides 284 through 306 of the UGT1A7 complementary DNA [cDNA]; GenBank accession number U85907) and antisense (1A7as1: 5′-GGTCCGCAAGATTGCAGTGGTTCG-3′; corresponding to nucleotides 1012 through 1032 of the UGT1A7 cDNA) primers homologous to UGT1A7 sequences within exons 1 and 3, respectively, were used in RT–PCR assays to identify the possible amplification of contaminating genomic DNA sequences. Reaction mixtures were incubated in a Perkin-Elmer 9700 thermocycler (Perkin-Elmer Corp., Foster City, CA) at 94 °C for 3 minutes, then for up to 41 amplification cycles at 94 °C for 1 minute, 64 °C for 1 minute, and 72 °C for 1 minute, followed by a final extension step for 7 minutes at 72 °C. Semiquantitative RT–PCR assays were performed on pooled tissue samples from each tissue site (n = 5 for tongue, n = 3 for floor of mouth, n = 3 for tonsil, n = 3 for larynx, n = 10 for esophagus, and n = 19 for liver) with the use of equimolar amounts of total RNA. Human β-actin exon 4 sense and exon 5 antisense primers [20 pmol for each primer; see (16) for primer sequences] were added after the ninth amplification cycle as an internal positive control for the presence and relative amount of RNA in each PCR assay. Reactions without RNA were included as negative controls in all RT–PCR experiments. Aliquots (10 μL) were removed from each PCR assay after 12, 35, 38, and 41 cycles of amplification, and the PCR products were resolved by electrophoresis in 8% polyacrylamide gels and visualized by staining the gels with 1 μg/mL ethidium bromide. Stained gels were photographed over UV light, and the DNA bands in the photograph that corresponded to PCR-amplified UGT1A7 and β-actin were scanned and analyzed with the use of a computerized densitometry program (Alpha Innotech Corp., San Leandro, CA). All of the RT–PCR bands were analyzed by using the entire lane width for the brightest band for any single gel, not including tailing, as the scan parameter. Densitometric readings of individual bands were plotted with the use of background corrections made for each individual gel by use of the same scanning parameters performed for the RT–PCR bands of the same gel. All of the RT–PCR analyses were performed in triplicate. The sequences of all of the PCR-amplified bands were confirmed by dye sequencing (17) (performed in the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center and Research Institute).

**Study Populations and Sample Processing**

Case subjects consisted of 137 patients diagnosed with primary squamous cell carcinoma of the oral cavity (which included cancer of the gingiva, hard palate, dorsal tongue, floor of mouth, inner lip, soft palate, buccal mucosa, tongue, tonsil, and oropharynx) and 57 patients diagnosed with primary squamous cell carcinoma of the larynx. Case subjects were identified and recruited during the period from 1994 through 2000 as part of a multi-institutional study conducted at Temple University Hospital (Philadelphia, PA), the Memorial Sloan-Kettering Cancer Center (New York, NY), and the New York Eye and Ear Infirmary (New York, NY). All cases of oropharyngeal cancer were diagnosed within 1 year of recruiting into the study and were confirmed histologically by the pathology departments at each of the respective institutions. Ninety-eight percent of the case subjects who were asked to participate in the study consented.

We recruited potential control subjects by screening the admission rosters of participating hospitals and soliciting outpatients who were visiting ear, nose, and throat or dental clinics for conditions unrelated to exposure to tobacco smoke, including cancer, heart disease (i.e., congestive heart failure, angina, coronary artery disease, myocardial infarction, and peripheral vascular disease), emphysema, or chronic obstructive pulmonary disease. Control subjects were recruited after an initial verbal screening to determine that they had no previous diagnosis of cancer. The eligible pool of control subjects was restricted to those individuals of the same age at diagnosis (±5 years), same race, and same sex as the case subjects; control subjects were matched in a 2-to-1 ratio to case subjects. Ninety-five percent of the control subjects who were asked to participate in the study consented.

To analyze the polymorphic genotypes of the subjects in our study, we collected buccal cell samples from 129 case subjects at a follow-up examination after surgery, or we used archived non-tumor oral tissue isolated from 65 case subjects at the time of surgery during tumor resection. Buccal cells were collected from 388 control subjects. Protocols involving the collection and analysis of buccal cell specimens were approved by the institutional review boards at each of the collaborating hospitals and were in accordance with assurances filed with and approved by the U.S. Department of Health and Human Services. Written informed consent was obtained from all of the study subjects.

A questionnaire that contained questions on demographics and life-long smoking habits was administered to all of the study subjects. Tobacco use was categorized at 94 °C for 1 minute, 64 °C for 1 minute, and 72 °C for 1 minute, followed by a final extension step for 7 minutes at 72 °C. Semiquantitative RT–PCR assays were performed on pooled tissue samples from each tissue site (n = 5 for tongue, n = 3 for floor of mouth, n = 3 for tonsil, n = 3 for larynx, n = 10 for esophagus, and n = 19 for liver) with the use of equimolar amounts of total RNA. Human β-actin exon 4 sense and exon 5 antisense primers [20 pmol for each primer; see (16) for primer sequences] were added after the ninth amplification cycle as an internal positive control for the presence and relative amount of RNA in each PCR assay. Reactions without RNA were included as negative controls in all RT–PCR experiments. Aliquots (10 μL) were removed from each PCR assay after 12, 35, 38, and 41 cycles of amplification, and the PCR products were resolved by electrophoresis in 8% polyacrylamide gels and visualized by staining the gels with 1 μg/mL ethidium bromide. Stained gels were photographed over UV light, and the DNA bands in the photograph that corresponded to PCR-amplified UGT1A7 and β-actin were scanned and analyzed with the use of a computerized densitometry program (Alpha Innotech Corp., San Leandro, CA). All of the RT–PCR bands were analyzed by using the entire lane width for the brightest band for any single gel, not including tailing, as the scan parameter. Densitometric readings of individual bands were plotted with the use of background corrections made for each individual gel by use of the same scanning parameters performed for the RT–PCR bands of the same gel. All of the RT–PCR analyses were performed in triplicate. The sequences of all of the PCR-amplified bands were confirmed by dye sequencing (17) (performed in the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center and Research Institute).

**Allelic discrimination of the codons 129/131 polymorphic site, we constructed control plasmid controls containing UGT1A7 exon 1 sequences homologous to either the wild-type or the polymorphic sequences at codons 129 and 131 to use as assay standards. Wild-type and polymorphic UGT1A7 exon 1 sequences were generated by PCR with the use of primers 1A7s1 and 1A7as2 (5′-TCCGGTACACAG-GGTTTGGAGAA-3′; corresponding to nucleotides 702 through 723 of the UGT1A7 cDNA) as the sense and antisense primers, respectively, and the buccal cell genomic DNA from a subject identified previously by dyeoxy DNA sequencing as a codon 129/131 heterozygote (i.e., an individual with one wild-type allele and one allele that is polymorphic at codons 129/131) as template. The resulting PCR products were cloned into plasmid pCR2.1 (InvitroGen Corp., Carlsbad, CA), and independent clones containing either wild-type UGT1A7 or the polymorphic 129/131 sequence were identified by dyeoxy sequencing (performed in the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center and Research Institute).

**Allelic discrimination of the codon 129/131 polymorphic site, the molecular core of the TaqMan Assay Reagents Allelic Discrimination protocol (Roche Pharmaceuticals, Branchburg, NJ), In
this assay, different probes are used in a PCR-based assay to distinguish between variant DNA sequences at a single locus. Briefly, 25-μL PCR reactions were performed in a 96-well plate with the use of (per well) 50 ng of buccal cell genomic DNA, 900 nM of the sense (1A7s2; 5’-GCACCATGCGAAGTGCA-T3’); corresponding to nucleotides 425 through 446 of the UGT1A7 cDNA) and antisense (1A7as3; 5’-GGATCGAGAAACACTGCACTCA-3’; corresponding to nucleotides 383 through 398 of the UGT1A7 cDNA) primers, 1× TaqMan Universal PCR Master Mix (Roche Pharmaceuticals), and 200 nM of the wild-type 6-carboxyfluorescein (FAM)-labeled and minor groove-binding fluorescent quencher (MBNQFQ)-conjugated probe (5’-FAM-TTAATGACCGAAAATT-MBNQFQ-3’) and the polymorphic VIC-labeled and MBNQFQ-conjugated probe (5’-VIC-TTTTAAGGA-CAAAAATT-MBNQFQ-3’) (Perkin-Elmer Corp.). The VIC-labeled probe was homologous to nucleotides 382 through 398 of the wild-type UGT1A7 cDNA; the FAM-labeled probe was homologous to nucleotides 383 through 398 of the wild-type UGT1A7 cDNA. For each 96-well plate, water was substituted for genomic DNA in four wells as a negative control for amplification, the wild-type UGT1A7 exon 1 cDNA was substituted for genomic DNA in eight wells (assay standard 1), and the polymorphic UGT1A7 exon 1 containing plasmid was substituted for genomic DNA in another eight wells (assay standard 2). PCR for allelic discrimination was performed for one cycle at 50°C for 2 minutes, one cycle at 95°C for 10 minutes, and 35 amplification cycles at 92°C for 15 seconds and 60°C for 1 minute in a GeneAmp 9700 thermocycler (Perkin-Elmer Corp.) with the use of the ABI Prism 7700 Sequence Detection System (version 1.7; Perkin-Elmer Corp.) within the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center and Research Institute. Study subjects with either the homozygous wild-type (129asn/131arg)/(129asn/131arg), homozygous polymorphic (129lys/131lys)/(129lys/131lys), or heterozygous (129asn/131lys)/(129lys/131arg) genotypes were identified by this analysis (Fig. 1, A), enabling us to discern the UGT1A7*1 and UGT1A7*4 alleles from the UGT1A7*2 and UGT1A7*3 alleles.

**Fig. 1.** Representative UGT1A7 genotyping assay results and expression of UGT1A7 transcripts in orolaryngeal tissue specimens. A) Computer-generated representation of UGT1A7 exon 129/131 polymorphism analysis. Samples with the homozgyous wild-type (129asn/131arg)/(129asn/131arg) genotype and the heterozygous (129asn/131lys)/(129lys/131lys) and homozygous polymorphic (129lys/131lys)/(129lys/131lys) genotypes are shown. The y- and x-axes show the relative fluorescence of FAM (for exon 129asn/131arg) or VIC (for exon 129lys/131lys), respectively, normalized with respect to each of the corresponding allelic controls used as references. Each plotted spot is indicative of the genotype for a single DNA specimen. B) Representative polymerase chain reaction (PCR)–restriction fragment length polymorphism (RFLP) analysis of the UGT1A7 exon 208 polymorphism. Lane M, DNA marker; lane 1, no template control; lane 2, undigested PCR-amplified UGT1A7 exon 1 fragment amplified from genomic DNA from a single subject; lanes 3 and 4, digested PCR-amplified UGT1A7 exon 1 fragment from two different subjects with the heterozygous wild-type UGT1A7 genotype; lanes 5 and 6, digested PCR-amplified UGT1A7 exon 1 fragment from two different subjects with the heterozygous UGT1A7 genotype; and lanes 7 and 8, digested PCR-amplified UGT1A7 exon 1 fragment from two different subjects with the homozygous polymorphic UGT1A7 genotype. DNA marker band sizes are indicated on left of panel; UGT1A7 PCR–RFLP band sizes are indicated on right of panel. C) Reverse transcription (RT)–PCR amplification of UGT1A7 transcripts from total RNA isolated from human orolaryngeal tissue specimens. Lane M, DNA marker; lane 1, no RNA control; lanes 2–6, total RNA from five different tongue specimens; lanes 7–9, total RNA from three different floor-of-mouth specimens; lanes 10 and 11, total RNA from two different tonsil specimens; lanes 12 and 13, total RNA from two different liver specimens; and lanes 14 and 15, total RNA from three different larynx specimens. DNA marker band sizes are indicated on left; UGT1A7 and β-actin RT–PCR band sizes are indicated on right. D) Levels of expression of UGT1A7 messenger RNA in human aerodigestive tract tissues as determined by semiquantitative RT–PCR with the use of β-actin as an internal control for expression levels. Expression of UGT1A7 and expression of β-actin were linear at all of the PCR cycles examined for this analysis. Error bars represent 95% confidence intervals. bp = base pairs.
The UGT1A7 codon 208 polymorphism was detected by PCR–RFLP analysis. Briefly, a 440-base-pair (bp) fragment was amplified by PCR in a 50-μL reaction volume that contained 50 ng of buccal cell genomic DNA, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM of each of deoxy-nucleotide triphosphates, 200 nM of the A17s1 and A17as2 primers, and 2.5 U of Taq DNA polymerase (Boehringer Mannheim). The reaction mixtures were incubated at 94 °C for 2 minutes, then for 41 amplification cycles at 94 °C for 30 seconds, at 57 °C for 30 seconds, and at 72 °C for 30 seconds, followed by a final extension step at 72 °C for 7 minutes. Ten microliters of each PCR sample was digested with 5 U of RsaI (New England Biolabs, Beverly, MA) at 37 °C for 1.5 hours and resolved on 8% native polyacrylamide gels to detect differences in RFLP patterns. The gels were stained with ethidium bromide and visualized by UV light. Three banding patterns were observed by RFLP analysis (Fig. 1, B): 1) a 440-bp band; 2) 337-bp and 103-bp bands; and 3) 189-bp and 135-bp bands that corresponded to the 208Arg/208Arg homozygous wild-type genotype (lanes 3 and 4); 2) 440-bp, 337-bp, and 103-bp bands that corresponded to the 208Arg/208Thr heterozygous genotype (lanes 5 and 6); and 3) 337-bp and 103-bp bands that corresponded to the 208Thr/208Thr homozygous polymorphic genotype (lanes 7 and 8). The presence of a RsaI site in codon 208 in the UGT1A7*3 and UGT1A7*4 alleles allowed us to distinguish the UGT1A7*1 and UGT1A7*2 alleles from the UGT1A7*3 and UGT1A7*4 alleles.

To confirm genotyping results, we performed the genotyping analysis twice for 40% of the specimens analyzed in this study. Selected PCR-amplified DNA samples from subjects possessing each of the nine detectable genotypes (Table 1) elucidated in this study were examined by dideoxy DNA sequencing (performed in the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center and Research Institute) to confirm UGT1A7 genotyping results.

**Statistical Analyses**

The risk of orolaryngeal cancer in relation to UGT1A7 genotypes was estimated with the use of unconditional logistic regression to calculate odds ratios (ORs) and 95% confidence intervals (CIs). For all of the analyses, the regression model included sex, age, race, region of subject recruitment (New York, NY, versus Philadelphia, PA), smoking (py), and alcohol consumption as categorical variables. Interactions between genotype and smoking or alcohol consumption were evaluated by stratifying subjects by the median number of py, 34 py, for all smokers or by categorical variables for alcohol consumption. In addition, the likelihood ratio test was used to compare the goodness of fit of the model with and without the various interaction terms. The chi-squared or Fisher’s exact tests were used, as appropriate, for analysis of allelic prevalence between racial groups and in case subjects versus control subjects. Two-tailed Student’s t test was used to compare continuous variables, such as age and cigarette consumption, between case and control subjects. Spearman’s nonparametric correlation test was used to examine the correlation between smoking and alcohol consumption. The statistical computer software SPSS (version 10.1) was used to perform all of the statistical analyses (20).

**RESULTS**

**Expression of UGT1A7 in Orolaryngeal Tissues**

To evaluate the potential physiologic importance of UGT1A7 in orolaryngeal tissue, we performed RT–PCR analysis to determine whether UGT1A7 is expressed in normal human tongue, tonsil, floor of mouth, and larynx. UGT1A7 was expressed in all orolaryngeal tissue specimens examined, but not in liver (Fig. 1, C). We compared the level of UGT1A7 messenger RNA (mRNA) with that of β-actin mRNA in the different tissues in a semiquantitative analysis of UGT1A7 expression. UGT1A7 expression and β-actin expression were linear at all of the cycles examined for this analysis. We found that the UGT1A7 mRNA level was highest in the tongue and lowest in the larynx (Fig. 1, D). We also found that UGT1A7 mRNA expression in orolaryngeal tissues was comparable to that observed for human esophagus.

**Cohort Characteristics**

A total of 125 Caucasian and 69 African-American subjects diagnosed with orolaryngeal cancer and 250 Caucasian and 138 African-American control subjects were entered in this study. Among the case subjects recruited from the collaborating institutions in New York City, 118 were Caucasian and 44 were African-American; among the control subjects recruited from the collaborating institutions in New York City, 125 were Caucasian and 56 were African-American. Among the case subjects recruited from Temple University Hospital in Philadelphia, seven were Caucasian and 25 were African-American; among the control subjects recruited from Temple University Hospital, 125 were Caucasian and 82 were African-American. The mean age of Caucasian and African-American case and control groups ranged from 59 to 61 years. Twenty-nine percent of the case and control subjects in each racial group were female. Approximately 87% of Caucasian case subjects and 58% of Caucasian control subjects were ever smokers (including cigar and pipe smokers), whereas 96% of African-American case subjects and 64% of African-American control subjects were ever smokers. The mean smoking dose (in py) for case and control subjects was 47 and 23, respectively, for Caucasians and 42 and 21, respectively, for African-Americans. Case subjects in both racial groups had a statistically significantly higher level (P<.001) of smoking than control subjects. A statistically significantly higher percentage of case subjects than control subjects were heavy drinkers for both African-Americans (46% of cases versus 17% of controls; P<.001) and Caucasians (34% of cases versus 13% of controls; P<.001). A statistically significant correlation (r = .43; P<.001) existed between smoking and alcohol consumption for all of the subjects in the entire cohort.

**Table 1. UGT1A7 allelic and polymorphism prevalence in orolaryngeal cancer case subjects and matched control subjects**

<table>
<thead>
<tr>
<th>Allele†</th>
<th>Polymorphism†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>129*omo131ns</td>
</tr>
<tr>
<td>Caucasians</td>
<td></td>
</tr>
<tr>
<td>Control subjects</td>
<td>206 (0.42)</td>
</tr>
<tr>
<td>Case subjects</td>
<td>88 (0.36)</td>
</tr>
<tr>
<td>African-Americans</td>
<td></td>
</tr>
<tr>
<td>Control subjects</td>
<td>103 (0.38)</td>
</tr>
<tr>
<td>Case subjects</td>
<td>46 (0.33)</td>
</tr>
</tbody>
</table>

†Numbers in parentheses refer to prevalence. Informative genotyping analysis was obtained for all subjects for the UGT1A7 codon 208 polymorphism but was not obtained for seven control subjects (five Caucasians and two African-Americans) and two case subjects (both Caucasian) for the UGT1A7 codon 129/131 polymorphism.

‡Case subjects had statistically significant increases in the prevalence of the combined UGT1A7*3 and UGT1A7*4 alleles (P = .01; 95% confidence interval = 0.35 to 0.52) or the codon 208*om polymorphism (P = .02; 95% confidence interval = 0.43 to 0.58) compared with control subjects.
UGT1A7 Alleles and Risk of Orolaryngeal Cancer

In a previous study, Guillemette et al. (5) described the prevalence of UGT1A7 alleles in a Caucasian cohort. In the present study, we have expanded on this study by comparing the prevalence of UGT1A7 alleles in Caucasians with that in African-Americans. We obtained informative PCR results concerning the identities of the polymorphisms at both UGT1A7 codons 129/131 and 208 for 123 and 245 of the Caucasian case and control subjects, respectively, and for 69 and 136 of the African-American case and control subjects, respectively. Among the Caucasian control subjects, the prevalence of the UGT1A7*1, UGT1A7*2, UGT1A7*3, and UGT1A7*4 alleles were 0.42, 0.24, 0.32, and 0.01, respectively (Table 1), and were similar to those reported in a previous study (5). The prevalence of individual UGT1A7 alleles was similar for control subjects who were recruited from institutions located in either New York City or Philadelphia for both racial groups (data not shown). The combined prevalence of the two low-activity alleles, UGT1A7*3 and UGT1A7*4, in African-American control subjects was 0.29 for those recruited from New York City and 0.30 for those recruited from Philadelphia. The combined prevalence of UGT1A7*3 and UGT1A7*4 in Caucasian control subjects from New York City and Philadelphia was 0.32 and 0.35, respectively. The combined prevalence of the UGT1A7*1 and UGT1A7*4 alleles was similar in control subjects of both racial groups. Caucasian control subjects had a statistically significantly (P < 0.001) lower frequency of the UGT1A7*2 allele (0.24) than African-American control subjects (0.39). In contrast, the frequency of the UGT1A7*3 allele was statistically significantly (P < 0.001) higher in Caucasian control subjects (0.32) than in African-American control subjects (0.23). The prevalence of the codon 129<sup>5</sup>/131<sup>B</sup> polymorphism in control subjects was similar for Caucasians (0.57) and African-Americans (0.61), whereas the prevalence of the codon 208<sup>A</sup>/B polymorphism in Caucasians (0.34) was statistically significantly higher (P = 0.003) than it was in African-Americans (0.23). The prevalence of both polymorphisms followed Hardy–Weinberg equilibrium.

To determine whether UGT1A7 alleles contribute to increased risk for orolaryngeal cancer, we examined the prevalence of UGT1A7 alleles in orolaryngeal cancer case subjects and compared it with that in control subjects for both Caucasians and African-Americans. Compared with Caucasian control subjects, Caucasian case subjects had a statistically significant increase (P = 0.01) in the prevalence of the combined low-activity UGT1A7 alleles (*3 and *4) and in the prevalence of the UGT1A7 codon 208<sup>B</sup> polymorphism (P = 0.02), which is present in both the UGT1A7*3 and UGT1A7*4 alleles (Table 1). A similar but statistically non-significant trend was observed for the prevalence of the combined low-activity UGT1A7 alleles (*3 and *4) as well as that of the UGT1A7 codon 208<sup>B</sup> polymorphism in African-Americans. We observed no statistically significant difference in the prevalence of the UGT1A7 codon 129<sup>5</sup>/131<sup>B</sup> polymorphism (present in both the UGT1A7*2 and UGT1A7*3 alleles) between case and control subjects in either racial group.

In a previous study, Guillemette et al. (5) showed that UGT1A7 isozymes encoded by variant UGT1A7 alleles exhibited substantially lower activity than the UGT1A7 isozyme encoded by the wild-type allele. We, therefore, stratified UGT1A7 genotypes into three categories based on their predicted phenotypes to determine whether UGT1A7 genotypes are correlated with risk for orolaryngeal cancer. UGT1A7*3/UGT1A7*3, UGT1A7*3/UGT1A7*4, and UGT1A7*4/UGT1A7*4 were considered to be low-activity genotypes; UGT1A7*1/UGT1A7*2, UGT1A7*1/UGT1A7*3, UGT1A7*1/UGT1A7*4, UGT1A7*2/UGT1A7*2, and UGT1A7*2/UGT1A7*3 were considered to be intermediate-activity genotypes; and UGT1A7*1/UGT1A7*1 was considered to be the high-activity (wild-type) genotype. We observed a statistically significant increase in risk for orolaryngeal cancer among all subjects who had the predicted low-activity UGT1A7 genotypes compared with subjects who had the predicted highest-activity genotype (OR = 3.7; 95% CI = 1.7 to 8.7; Table 2). Statistically significant increases in risk for orolaryngeal cancer were observed for both Caucasians (OR = 2.8; 95% CI = 1.1 to 7.6) and for African-Americans (OR = 6.2; 95% CI = 1.2 to 31) who had the predicted low-activity UGT1A7 genotypes. This increase in risk was also observed when we compared all of the subjects who had the predicted low-activity UGT1A7 genotypes with (a) subjects who had the predicted high-activity genotype (UGT1A7*1/UGT1A7*1) or the predicted highest activity of the intermediate-activity genotypes (UGT1A7*1/UGT1A7*2; OR = 3.1; 95% CI = 1.6 to 6.0), (b) subjects who had either the predicted high-activity genotype (UGT1A7*1/UGT1A7*1) or any of the predicted intermediate-activity genotypes (UGT1A7*1/UGT1A7*2, UGT1A7*1/UGT1A7*3, UGT1A7*1/UGT1A7*4, UGT1A7*2/UGT1A7*2, and UGT1A7*2/UGT1A7*3; OR = 2.8; 95% CI = 1.5 to 5.2), or (c) subjects who had any of the predicted intermediate-activity genotypes (OR = 2.6; 95% CI = 1.4 to 4.9) (data not shown). For subjects with the predicted low-activity genotypes, the risks of oral cavity cancer (OR = 4.2; 95% CI = 1.7 to 10) and laryngeal cancer (OR = 3.7; 95% CI = 0.99 to 14) were similar (Table 2). A statistically significant increase in orolaryngeal cancer risk was also observed for the UGT1A7*3/UGT1A7*3 genotype (OR = 3.9; 95% CI = 1.6 to 9.2) when analyzed individually when compared with the predicted high-activity genotype (UGT1A7*1/UGT1A7*1) (data not shown). Analysis of other predicted high-activity genotypes could not be performed adequately because of the low prevalence of the UGT1A7*4 allele in the population.

To examine the relationship between the UGT1A7 genotype and the risk of orolaryngeal cancer by exposure to environmental risk factors, we stratified study subjects by the enzymatic activity predicted for their UGT1A7 genotypes and either smoking history or alcohol consumption (Table 3). To obtain sufficient power for this analysis, we considered subjects who had either the UGT1A7*1/UGT1A7*1 or the UGT1A7*1/UGT1A7*2 genotypes as having predicted high-activity UGT1A7 genotypes because we found that, as described above, the use of this genotype combination as the reference group resulted in similar risk for predicted low-activity UGT1A7 genotypes as when only the UGT1A7*1/UGT1A7*1 genotype was the reference group. We observed no statistically significant association between the UGT1A7 genotype and orolaryngeal cancer risk in never smokers. In contrast, study subjects with predicted low-activity UGT1A7 genotypes who were classified as either light or heavy smokers had statistically significant increases in orolaryngeal cancer risk.
Risk for orolaryngeal cancer was 8.5- to 10-fold higher for subjects with predicted low-activity UGT1A7 genotypes than for subjects with predicted high-activity genotypes within each smoking group (OR = 3.7 [95% CI = 1.1 to 12] for light smokers; OR = 6.1 [95% CI = 1.5 to 25] for heavy smokers; data not shown). However, a formal test for interaction between the UGT1A7 genotype and the smoking dose was not statistically significant in this analysis (P = .37 by the likelihood ratio test).

In contrast to the differences in orolaryngeal cancer risk observed at different smoking exposures, predicted low-activity UGT1A7 genotypes were statistically significantly associated with increased risk for orolaryngeal cancer regardless of the alcohol-consumption category (Table 3). A formal test for interaction between UGT1A7 genotype and alcohol consumption was not statistically significant in this analysis (P = .67 by the likelihood ratio test). As expected, we observed a statistically significant increase in risk for orolaryngeal cancer for all subjects, regardless of UGT1A7 genotype, as their exposure to tobacco smoke or their consumption of alcohol increased.

### Table 2. Predicted low- and intermediate-activity versus high-activity UGT1A7 genotypes and orolaryngeal cancer risk*

| Predicted activity of UGT1A7 genotypes† | Total cohort: OR (95% CI)‡,§ | Caucasian: OR (95% CI)¶,|| | African-American: OR (95% CI)¶,# | Oral cavity**: OR (95% CI)¶,§ | Larynx††: OR (95% CI)¶,†‡ |
|----------------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| High                                   | 1.0 (referent)              | 1.0 (referent)              | 1.0 (referent)              | 1.0 (referent)              | 1.0 (referent)              | 1.0 (referent)              |
| Intermediate                           | 1.5 (0.78 to 2.7)           | 1.1 (0.49 to 2.3)           | 2.2 (0.73 to 6.5)           | 1.4 (0.70 to 2.9)           | 1.7 (0.62 to 4.6)           |
| Low                                    | 3.7 (1.7 to 8.7)            | 2.8 (1.1 to 7.6)            | 6.2 (1.2 to 31)             | 4.2 (1.7 to 10)             | 3.7 (0.99 to 14)            |

*OR = odds ratio; CI = confidence interval.
†Genotypes were stratified according to predicted activities based on allele-encoded isozyme glucuronidating activities toward benzo[a]pyrene metabolites (7).
‡ORs (95% CIs) were calculated after adjustment for race, sex, age, region of subject recruitment, and smoking (in pack-years).
§Eighteen subjects were excluded from this analysis because of incomplete genotyping (n = 9) or incomplete alcohol consumption or smoking data (n = 9).
¶OR and 95% CI calculated after adjustment for sex, age, region of subject recruitment, smoking (in pack-years), and alcohol consumption (categorical variables).
**Analysis included subjects with oral cavity cancer versus total control subjects.
††Analysis included subjects with laryngeal cancer versus total control subjects.
‡‡Analysis did not include 14 subjects with incomplete genotyping (n = 7) or alcohol consumption or smoking (n = 7) data.
To our knowledge, this is the first study to examine the association between polymorphic variants of UGTs and the risk for a tobacco-related cancer. Consistent with the observed lower activity of the UGT1A7 isozymes encoded by the UGT1A7*3 and UGT1A7*4 alleles against several BaP metabolites (5), we found that individuals with genotypes containing different combinations of these two alleles had an increased risk for oropharyngeal cancer. Moreover, these predicted low-activity UGT1A7 genotypes were associated with an increased risk for oropharyngeal cancer in both Caucasians and African-Americans, and a similar risk was observed for both cancer of the oral cavity and cancer of the larynx. Although we found that the interaction between the UGT1A7 genotype and smoking was not statistically significant, the association between UGT1A7 genotype and oropharyngeal cancer risk was linked to exposure to tobacco smoke on the basis of the fact that a statistically significant risk was observed in smokers but not in never smokers. In addition, RT–PCR analysis of normal human tissues demonstrated that UGT1A7 is expressed in oropharyngeal tissues at levels similar to those observed for tissue sites such as the esophagus where the UGT1A7 expression is high (6). This observation is consistent with the finding that tissue microsomes from all upper aerodigestive tract sites tested in our laboratory exhibit substantial activity against metabolites of BaP (Fang JL, Zheng Z, Lazarus P: unpublished results) as well as with the previously described role for UGT1A7 in the metabolism of several tobacco carcinogens (3–5). Considered together, these data suggest that UGT1A7 plays an important role in the detoxification of tobacco carcinogens within the upper aerodigestive tract and that genetic alterations that alter UGT1A7 activity may affect smoking-related cancer risk.

We are aware of two potential limitations of this study. First, we could not match control and case subjects by the institution or geographic region of case subject recruitment because we were unable to identify and recruit suitable hospital-based patients who fit our exclusion criteria for control subjects from a collaborating cancer center. However, the prevalences of both UGT1A7 alleles and polymorphisms were highly similar for control subjects recruited from other collaborating institutions in New York and Philadelphia for both Caucasians and African-Americans. In addition, we found that the association between low-activity UGT1A7 genotypes and risk of orolaryngeal cancer was statistically significant regardless of whether the data were analyzed with or without adjustment for region or institution of subject recruitment. This observation suggests that failure to match control subjects with case subjects by region or institution of case recruitment was not a confounder in these studies.

A second potential limitation of this study was the fact that the control subjects who were recruited into this study were hospital based; therefore, they may not reflect the overall prevalence of UGT1A7 alleles in the general population. However, as part of the International Project on Genetic Susceptibility to Environmental Carcinogens database, we performed a large meta-analysis of more than 15 000 individuals without cancer and found no statistically significant differences in the overall prevalence of metabolizing enzyme polymorphisms between hospital-based and population-based control subjects (21). Therefore, it is unlikely that the use of a hospital-based cohort for control subject recruitment plays a confounding role in the statistically significant associations between UGT1A7 genotype and oropharyngeal cancer risk that we observed in this study.

It remains to be determined whether the UGT1A7 genotype plays a role in the susceptibility to other tobacco-related cancers, including cancers of the lung or esophagus. Although studies in our laboratory have confirmed the results obtained by Strassburg et al. (6), who demonstrated that normal human esophageal tissue expresses UGT1A7, we have not yet been able to demonstrate that UGT1A7 is expressed in normal human lung tissue (Zheng Z, Lazarus P: unpublished data). Because UGT1A7 is an extrahepatic enzyme, the effects of the UGT1A7 genotype on cancer risk are likely to depend on the expression levels of UGT1A7 in the target tissues, where the mutagenic effects of carcinogen metabolism are manifested. Further studies examining UGT1A7 genotype will, therefore, be required to elucidate the role of UGT1A7 in cancer risk in other potential target tissues.

REFERENCES

(12) Coffman BL, King CD, Rios GR, Tophily TR. The glucuronidation of opioids, other xenobiotics, and androgens by human UGT2B7(268) and UGT2B7H(268). Drug Metab Dispos 1998;26:733–7.
(16) Strassburg CP, Oldhafer K, Manns MP, Tukey RH. Differential expression of the UGT1A lo-


NOTES

Supported by Public Health Service (PHS) grants DE12206 and DE13158 (National Institute of Dental Research and Craniofacial Research) to P. Lazarus and PHS grant CA68384 (National Cancer Institute) (P. Lazarus, project leader; Steven Stellman, principal investigator) from the National Institutes of Health, Department of Health and Human Services.

We are grateful to Mary Beth Colter and Lanmin Zhang, of the Molecular Biology Core Facility, and Marek Wloch, of the Tissue Procurement Facility, at the H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, for their help and excellent technical assistance.

Manuscript received December 7, 2000; revised July 11, 2001; accepted July 27, 2001.