An Association Between Glutathione S-Transferase P1 Gene Polymorphism and Younger Age at Onset of Lung Carcinoma

David P. Miller, ScD[1]
Kofi Asomaning, MBChB, MS[1]
Geoffrey Liu, MD[1]
John C. Wain, MD[2]
Thomas J. Lynch, MD[3]
Donna Neuberg, ScD[4]
Li Su, BS[1]
David C. Christiani, MD[1,5]

[1] Department of Environmental Heath, Occupational Health Program, Harvard School of Public Health, Boston, Massachusetts.

[2] Thoracic Surgery Unit, Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.

[3] Oncology Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.

[4] Department of Biostatistics, Harvard School of Public Health, Dana-Farber Cancer Institute, Boston, Massachusetts.

[5] Pulmonary and Critical Care Unit, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.

Supported by The National Institutes of Health. (Grants CA 74386, ES/CA 06409, and ES00002).

Dr. Miller was supported by a Flight Attendants Medical Research Institute Young Clinical Scientist Award.

We gratefully acknowledge the assistance of Linda Lineback, Barbara Bean, Andrea Shafer, SooAe Shaneyfelt, Jessica Shin, Jeanne Jackson, and Andrea Solomon for patient recruitment and data collection; Lucy Ann Principe, Salvatore V. Mucci, and Richard Rivera-Massa for data entry; Stephanie Shih for sample preparation and genotyping; and Geoffrey Liu and Shannon Magari for editing advice and comments.

BACKGROUND. Among the genes that encode the glutathione S-transferase (GST) superfamily of Phase 2 metabolizing enzymes, *GSTP1* has the highest expression in the lung. The polymorphic *GSTP1* gene encodes glutathione S-transferase π, which is an enzyme that detoxifies cigarette carcinogens, such as benzo-[a]-pyrene. The variant *GSTP1 GG* genotype is associated with lower enzymatic activity and higher DNA adduct levels in human lymphocytes compared with the *AA* genotype.

METHODS. The authors evaluated the association of *GSTP1* genotypes with lung cancer in 1921 cases and 1343 controls of Caucasian descent by using polymerase chain reaction-restriction fragment length polymorphism techniques. The results were analyzed with multiple logistic regression adjusting for age, gender, smoking status, and pack-years. To investigate specifically the subset of younger lung cancer patients and controls, the effect of age (either as a dichotomous or continuous variable in separate models) was analyzed as a modifying factor of the association between the *GSTP1* polymorphism and lung cancer.

RESULTS. The *GSTP1 GG* genotype was not associated with an overall increased risk of lung cancer (adjusted odds ratio, 1.02; 95% confidence interval [95% CI], 0.78–1.34) compared with the *GSTP1 AA* genotype. In both models that evaluated the gene-age interaction, an overall statistically significant interaction ($P < .01$) was observed between age and the *GG* genotype. However, for the model that included age as a dichotomous variable, the odds ratio of lung cancer risk with the *GG* genotype compared with the *AA* among individuals age $\leq$50 years was 2.67 (95% CI, 1.36–5.22); in older individuals, the risk was 0.87 (95% CI, 0.65–1.2).

CONCLUSIONS. The *GSTP1 GG* genotype was associated with increased lung cancer susceptibility among younger study participants. ***Cancer*** **2006;107:1570–7.**

KEYWORDS: *GSTP1*, early age onset lung carcinoma, case control, logistic regression, interaction.

Young individuals with lung cancer may be an optimal subset of patients in whom to study disease susceptibility genes. In older individuals, the cumulative effect of smoking may mask underlying genetic predispositions, because smoking has such a high attributa-

Address for reprints: David C. Christiani, MD, Harvard School of Public Health, 665 Huntington Avenue, Building 1, Room 1402, Boston, MA 02115; Fax: (617) 432-3441; E-mail: dchristi@hsph.harvard.edu

Received March 22, 2006; revision received May 22, 2006; accepted June 9, 2006.

DOI 10.1002/cncr.22124
Published online 24 August 2006 in Wiley InterScience (www.interscience.wiley.com).

ble risk in lung cancer. Kreuzer et al. reported that having a smoking first-degree relative with early-onset carcinoma may increase the risk of early-onset lung carcinoma (odds ratio [OR], 5.65; 95% confidence interval [95% CI], 0.7–46.9),[1] although their results were not statistically significant. To our knowledge to date, early-onset lung cancer has been associated with a specific lung cancer subtype (adenocarcinomas),[2,3] although the effect of genetic polymorphisms on early-onset lung carcinoma has not been studied extensively. Cote et al. and Taioli et al.[4,5] described the association between metabolizing genotypes and early-onset lung cancer, but those investigators did not assess the effect modification of age at onset. Early-onset lung carcinoma has been associated with specific variant genotypes, which have been related to reduced function.

Certain genetic polymorphisms of metabolizing genes have been associated with altering lung cancer risk by modifying the effect of tobacco smoke carcinogens in the lung. The glutathione S-transferase (*GST*) *M1* gene (*GSTM1*) polymorphism is an example. Its association with lung cancer as an independent risk factor and its role in modifying smoke effects on lung cancer risk have been studied extensively.[6–9] London et al.[9] and Saarikoski et al.[10] suggested that the effect of the *GSTM1* null genotype is statistically significant only in populations in which cumulative pack-years are relatively low (<40 pack-years). In the lung, the *GSTP1* gene has the highest expression of all genes in the GST superfamily. In addition, the *GSTP1* polymorphism codes for an enzyme, GST class π, which is a major metabolizer of the activated products of benzopyrene (B[a]P), a principal carcinogen in cigarette smoke.[11–13] Thus, its function provides a strong rationale for studying this polymorphism in relation to lung cancer risk.

A functional genetic polymorphism of *GSTP1* is the result of a single base-pair (bp) substitution in which (A) adenine is replaced by (G) guanine, leading to an amino acid substitution in which isoleucine is replaced by valine. This substitution results in a disabled detoxification function and, thus, the accumulation of higher levels of activated carcinogens in the lung. The *GSTP1 GG* homozygous genotype is associated with higher adduct levels in human lymphocytes[14] and significantly lower enzymatic activity.[15] Recent studies[7,9,16–19] that examined the association between *GSTP1* polymorphisms and lung cancer reported no statistically significant associations. Because sample sizes for these studies ranged from 169 to 639 patients, larger studies may help clarify some of those findings.

After adjusting for exposure to smoking, carrying the *GSTP1 GG* genotype may lead to an earlier onset of lung cancer because of prolonged exposure to tobacco carcinogens, even in younger individuals. We tested this hypothesis and investigated the role of *GSTP1 GG* as an independent risk factor in a large case–control study.

MATERIALS AND METHODS

Participants in the current study were recruited as part of an ongoing case–control study that was initiated in 1992 at the Massachusetts General Hospital (MGH), in Boston. The study was approved by the Institutional Review Boards at both MGH and the Harvard School of Public Health. Eligible cases included any individual age $\geq$18 years with a diagnosis of primary lung cancer who was evaluated by the Pulmonary, Thoracic, or Hematology-Oncology Unit at MGH for either for surgery (from 1992) or for chemotherapy, radiation treatment, or any combination of treatment modalities (from 1996). An MGH lung pathologist histologically confirmed all diagnoses. Controls were recruited from among friends and nonblood-related family members of patients in the case group (usually spouses) (41%). If friends of lung cancer patients were not available, then controls were recruited from among friends and family of patients who either were undergoing thoracic surgery, receiving chemotherapy, or receiving radiation treatment for a condition other than lung cancer (59%). To reduce potential variation in allele frequency by ethnicity, only white patients (97% of our genotyped study population) for whom there was complete information regarding gender, age, and smoking variables (smoking status, pack-years) were analyzed. Our objective was to select a large, unmatched set of controls with a broad spectrum of variability in age, gender, and smoking variables to allow for effect modification evaluation of these variables. To determine whether our set of controls were similar with regard to average smoking habits to the general Massachusetts population, we compared smoking exposure covariate data obtained from our controls with information provided by the Massachusetts Tobacco Survey (1993–1997)[20] and observed that they were similar. The variables that were available for comparison were mean cigarettes per day, smoking status, age started smoking, and, for exsmokers, the number of years quit smoking. The participation rate was >85% and did not differ between cases and controls. Although recruitment was not restricted to residents of Massachusetts, 84% of cases and 79% of controls were from Massachusetts, and

another 10% of cases and 11.5% of controls were from the New England area.

Blood samples were collected from all participants at the time of recruitment. Two to three 10-mL ethylenediamine tetracetic acid tubes and a single 5-mL silicon-coated tube were used for sample collection. Samples were processed in the Molecular Epidemiology Laboratory at the Harvard School of Public Health. DNA was extracted from whole blood for the purpose of genotyping. *GSTP1* genotypes were determined by using polymerase chain reaction-restriction fragment length polymorphism methods, as described by Harris et al.,[17] without the knowledge of case or control status. The presence of a 330- base pair (bp) fragment identified the I_{105} allele; the V_{105} allele was detected by the presence of a 220-bp fragment and a 110-bp fragment; the heterozygote possessed all 3 fragments. For quality control, a random 5% of the samples were repeated with 100% compliance. Two authors independently reviewed 100% of the agarose gels.

Two research nurses directly administered a health questionnaire. Some participants opted to complete the questionnaire at home and return it by mail in a self-addressed, stamped envelope. Participants were contacted by telephone when there were missing data. Other variables were obtained through the health questionnaire.[6,21,22] Age, gender, race, weight, education, medical history, smoking history, family history of cancer, work history, exposure to various substances, participation in many activities, and food preparation and consumption data were collected. Smoking status was defined as nonsmoker (smoked <1 cigarette per day for <1 year), exsmoker (at the time of diagnosis) and current smoker, (at the time of diagnosis). Pack-years were calculated to estimate the cumulative exposure to smoking by multiplying the number of packs smoked per day by the number of years smoked.

Population characteristics were tabulated, and significant differences in the distribution of the principal covariates were tested by using the chi-square, Fisher exact, and Student *t* tests, as appropriate. Genotype frequencies among controls were tested for conformity to Hardy–Weinberg equilibrium. Multiple logistic regression was used to assess the association between *GSTP1 GG* and lung cancer risk, adjusting for age, gender, indicator variables for smoking status (nonsmoker, exsmoker, and current smoker), a continuous variable for cumulative smoking exposure (pack-years), and an indicator variable for the heterozygote. To evaluate whether *GSTP1* is associated with early-onset lung carcinoma, we considered age as an effect modifier of the association between *GSTP1* and lung cancer risk. We evaluated the interaction between genotype and age, in which age was considered either as a continuous variable or as a dichotomous variable. When age was included as a dichotomized variable (age $\leq$50 years and age $>$50 years), the interaction term was the product of this variable and the indicator variable for *GSTP1 GG*. Because the average age of lung cancer patients in our sample was 66 years, using a cutoff age of 50 years appropriately included all younger patients in the evaluation. To determine whether the interaction terms were significant, we used a likelihood-ratio test to compare the 2 nested models[23] with and without the interaction term. All analyses were adjusted for smoking status, gender, and pack-years. Several analyses were conducted to determine whether disease stage or cumulative smoking dose (pack-years) may confound the association between age at diagnosis and genotype. To determine whether disease stage was a factor that contributed to the results, the cases were stratified into 2 groups: patients with early-stage disease and patients with late-stage disease. Two multiple logistic regression analyses were performed comparing the control group with both case groups individually. To evaluate the role of pack-years, the main multiple logistic regression analysis was stratified into groups divided by the median pack-years among the control.[24] The SAS statistical package[25] was used to perform all analyses.

RESULTS

Table 1 shows the distribution of demographic, histologic, and genotypic characteristics among cases and controls age $\leq$50 years and those age $>$50 years. The distributions of the covariates described above within age categories were similar. The distributions of the genotypes in the control group for all of the strata considered were in Hardy–Weinberg equilibrium. Women were represented equally among the cases (50.3% and 52.2%, respectively); they were under represented among controls (41.9% and 45.8, respectively). The frequency of current smokers was greater among cases than among controls in both age categories. The frequency of exsmokers was slightly greater among cases, whereas the frequency of nonsmokers and exsmokers was significantly higher among controls ($P < .0001$). Education level was not available for 10% of study participants (322 cases, 34 controls). Family history of cancer was more frequent among cases ($P \leq .02$). Almost 50% of the reported cancer cell types were adenocarcinoma. The relatively high percentage of early-stage (potentially curable) cancers in our sample reflected a referral bias

TABLE 1
Descriptive Characteristics of the Study Population Stratified by Age Category and Case Status*

	No. of patients (%)					
	Age ≤50 years			Age >50 years		
Characteristic	Cases (*n* = 159)	Controls (*n* = 351)	*P*	Cases (*n* = 1762)	Controls (*n* = 992)	*P*
Median age (range), y	45.1 (30.4–50.0)	43.5 (26.7–50.0)	.0005	67.9 (50.1–90.7)	64.2 (50.1–96.3)	<.0001
Gender						
Male	79 (49.7)	204 (58.1)	.10	842 (47.8)	537 (54.1)	.001
Female	80 (50.3)	147 (41.9)		920 (52.2)	455 (45.8)	
Smoking history						
Nonsmokers	21 (13.2)	152 (43.3)	.0001	122 (6.9)	325 (32.8)	<.0001
Exsmokers	37 (23.3)	101 (28.8)		983 (55.8)	510 (51.4)	
Current smokers	101 (63.5)	98 (27.9)		657 (37.3)	157 (15.8)	
Ever smokers: Median pack-years (range)	31.1 (0.2–147.8)	18.0 (0.1–100.6)	<.0001	54.0 (0.03–231.0)	28.5 (0.03–210.0)	<.0001
College degree	49 (30.8)	128 (36.5)	.003	388 (22.0)	282 (28.4)	<.0001
No college degree	95 (59.8)	214 (61.0)		1067 (60.6)	685 (69.1)	
Not available	15 (9.4)	9 (2.5)		307 (17.4)	25 (2.5)	
Family history of lung cancer			.02			.002
Yes	27 (17.0)	31 (8.8)		311 (17.6)	127 (12.8)	
No	119 (74.8)	297 (84.6)		1311 (74.4)	796 (80.2)	
Not available	13 (8.2)	23 (6.6)		140 (8.0)	69 (7.0)	
Cell type						
Adenocarcinoma	85 (53.5)			754 (42.8)		
Squamous cell	23 (14.5)			389 (22.1)		
Large cell	16 (10.1)			130 (7.4)		
Small cell	7 (4.4)			169 (9.6)		
Bronchioalveolar	9 (5.7)			166 (9.4)		
Mixed	6 (3.8)			48 (2.7)		
>1 primary	5 (3.1)			26 (1.5)		
Nonsmall cell not otherwise classified	8 (3.6)			62 (3.5)		
Not available	5 (3.1)			18 (1.0)		
Stage						
Early stage/limited disease	46 (28.9)			919 (52.2)		
Late stage/extensive disease	113 (71.1)			843 (47.8)		

* Medians were tested with the Wilcoxon rank-sum test, and frequency was tested with the chi-square test.

because of the surgical expertise at MGH, which was the institution of preference for surgical patients in the earlier years of recruitment (1992–1996). We did not have an a priori hypothesis leading us to believe that stage would have an impact on our results.

The distribution of the covariates from Table 1, stratified by the different polymorphisms, is shown in Table 2. The *GSTP1 GG* and *AA* genotypes were slightly more frequent among cases than controls in both age categories; conversely, heterozygote status was proportionally lower among cases than controls. Cases with *GSTP1 GG* were slightly younger (median age, 66.4 years) than the other cases, whereas controls with *GSTP1 GG* were slightly older than other controls (median age, 61.0 years). Among patients in the case group who had the *GSTP1 GG* polymorphism, 13% were age <50 years compared with the control group (18% of controls were age <50 years). The frequency of the *GSTP1 GG* genotype was higher among men than among women in cases (56% vs. 10%, respectively) and did not differ among controls (50% and 50%). The distribution of the heterozygote *AG* was higher among men than among women for cases (52% and 48%, respectively), but the reverse was true in controls (43%, 57%). The distribution of genotypes across different educational levels did not differ. The frequency of the *GSTP1 GG* genotype was 43% for patients in the case group who had adenocarcinoma cell histology and 22% for patients in the case group who had squamous cell histology.

Among cases, the frequencies of the *GSTP1 AA*, *AG*, and *GG* genotypes were greatest in current smokers and exsmokers and lowest in nonsmokers. For the control group, the lowest frequency for all genotypes was in the current smokers category.

Table 3 summarizes the results from the multiple logistic regression models on all cases and controls. The overall effect of *GSTP1* polymorphism on lung

TABLE 2
Descriptive Characteristics of the Study Population Stratified by Case Status and Genotype*

	No. of patients (%)							
	Cases				Controls			
Characteristic	*GSTP1 AA* (*n* = 885)	*GSTP1 AG* (*n* = 816)	*GSTP1 GG* (*n* = 220)	*P*	*GSTP1 AA* (*n* = 579)	*GSTP1 AG* (*n* = 623)	*GSTP1 GG* (*n* = 141)	*P*
Age								
Median (range), y	66.9 (30.7–88.8)	66.7 (30.4–90.7)	66.4 (36.2–87.8)	.47	60.7 (27.1–83.9)	58.3 (26.7–96.3)	61.0 (34.3–83.8)	.007
>50	822 (92.9)	748 (91.7)	192 (87.3)	.03	436 (75.3)	441 (70.8)	115 (81.6)	.02
≤50	63 (7.1)	68 (8.3)	28 (12.7)		143 (24.7)	182 (29.2)	26 (18.4)	
Gender								
Male	448 (50.6)	427 (52.3)	124 (56.4)	.30	264 (45.6)	268 (43.0)	70 (49.6)	.32
Female	437 (49.4)	389 (47.7)	96 (43.6)		315 (54.4)	355 (57.0)	71 (50.4)	
Smoking history								
Nonsmokers	64 (7.2)	66 (8.1)	13 (5.9)	.81	199 (34.4)	230 (36.9)	48 (34.1)	.74
Exsmokers	267 (52.8)	431 (52.8)	122 (55.5)		272 (47.0)	271 (43.5)	68 (48.2)	
Current smokers	354 (40.0)	319 (39.1)	85 (38.6)		108 (18.6)	122 (19.6)	25 (17.7)	
Ever smokers: Median pack-years (range)	52.0 (0.2–231.0)	51.3 (0.03–209)	52.0 (0.2–186.0)	.89	24.0 (0.03–210.0)	26.5 (0.1–152.0)	26.5 (0.05–100.6)	.85
Education level								
College degree	197 (22.3)	181 (22.2)	59 (26.8)	.20	171 (29.5)	197 (31.6)	42 (29.8)	.57
No college degree	547 (61.8)	483 (59.2)	132 (60.0)		396 (68.4)	410 (65.8)	93 (66.0)	
Not available	141 (15.9)	152 (18.6)	29 (13.2)		12 (2.1)	16 (2.6)	6 (4.2)	
Family history of lung cancer				.39				.95
Yes	161 (18.2)	146 (17.9)	31 (14.1)		71 (12.3)	72 (11.6)	15 (10.6)	
No	660 (74.6)	604 (74.0)	166 (75.5)		466 (80.5)	511 (82.0)	116 (82.3)	
Not available	64 (7.2)	66 (8.1)	23 (10.4)		42 (7.3)	40 (6.4)	10 (7.1)	
Cell type								
Adenocarcinoma	402 (45.4)	343(42.0)	94 (42.7)					
Squamous cell	190 (21.5)	173 (21.2)	49 (22.3)					
Large cell	69 (7.8)	62 (7.6)	10 (6.8)					
Small cell	69 (7.8)	80 (9.8)	27 (12.3)					
Bronchioalveolar	79 (8.9)	77 (9.4)	19 (8.6)					
Mixed	25 (2.8)	26 (3.2)	3 (1.4)					
>1 Primary tumor	10 (0.9)	14 (1.7)	2 (1.1)					
Nonsmall cell not otherwise classified	31 (3.6)	31 (3.8)	8 (3.5)					
Not available	10 (1.1)	10 (1.2)	3 (1.4)					
Stage								
Limited	438 (49.5)	405 (49.6)	122 (55.4)					
Extensive	447 (50.5)	411 (50.4)	98 (44.6)					

GSTP1 indicates glutathione S-transferase P1.

* Medians were tested by using the Wilcoxon rank-sum test, and frequency was tested by using the chi-square test.

cancer risk was assessed, and the adjusted OR for lung cancer associated with the *GSTP1 GG* genotype, compared with the *GSTP1 AA* genotype, was 1.02 (95% CI, 0.78–1.34). We observed an increased risk of lung cancer among younger participants. The interaction terms between the *GSTP1 GG* polymorphism and age for both models (a model with age as a dichotomous variable and a model with age as a continuous variable) was statistically significant ($P < .01$). The trend for both models was the same and showed the same greater association between *GSTP1 GG* and lung cancer with younger age. The adjusted OR of lung cancer associated with *GSTP1 GG* compared with *GSTP1 AA* among individuals age ≤50 years was 2.67 (95% CI, 1.36–5.22). Among older individuals, the OR was 0.87 (95% CI, 0.65–1.2). None of the interaction terms between *GSTP1 AG* and age were statistically significant. In comparing the patients who had early-stage disease in the case group with all controls or the patients who had late-stage disease in the case group with all controls, no differences were observed in the *GSTP1 GG*-age interaction or in the association with lung cancer risk. In addition, analyses that were stratified by the median pack-years in the control group did not alter the *GSTP1 GG*-age

TABLE 3
Multiple Logistic Regression Results

Analysis	No. of patients	OR	95% CI	*P*
Overall effect				
Crude				
AA	1464	1.00 (Referent group)		
AG	1439	0.86	0.74–0.99	.04
GG	361	1.02	0.81–1.29	.86
Adjusted*				
AA	1464	1.00 (Referent group)		
AG	1439	0.90	0.76–1.34	.23
GG	361	1.02	0.78–1.34	.89
Among patients age ≤50 y				
Crude				
AA	206	1.00 (Referent group)		
AG	250	0.85	0.57–1.27	.42
GG	54	2.44	1.33–4.50	.004
Adjusted*				
AA	206	1.00 (Referent group)		
AG	250	0.89	0.57–1.38	.60
GG	54	2.67	1.36–5.22	.004
Among patients age >50 y				
Crude				
AA	1258	1.00 (Referent group)		
AG	1189	0.90	0.76–1.06	.21
GG	307	0.89	0.69–1.15	.36
Adjusted*				
AA	1258	1.00 (Referent group)		
AG	1189	0.89	0.74–1.07	.20
GG	307	0.87	0.65–1.16	.35

OR indicates odds ratio; 95% CI, 95% confidence interval.

* Analyses were adjusted for age, pack-years, smoking status (indicator variables for exsmokers and current smokers), gender, and an indicator variable for the heterozygote. Interaction terms for GG were significant both for crude and adjusted analyses and for dichotomous and continuous variables. The intercation term for AG was not significant for any of the analyses.

interaction or the association with lung cancer risk (data not shown).

DISCUSSION

Previous studies have indicated the possibility of a gene effect in early-onset lung cancer. Bailey-Wilson et al. reported the location of a gene locus that, controlling for smoke exposure and acting in concert with it, may predispose individuals to earlier onset lung cancer.[26] Gauderman and Morrison also reported the presence of a susceptibility locus that may lead to early-onset lung cancer.[27] Yang et al. observed that the high risk of a gene that contributes to early-onset lung cancer was more evident in nonsmokers and decreased with age as the effect of smoking increased with age.[24] At older ages, patients also may die of other causes related to their smoking exposure, reducing the overall contribution of the gene.[26] Genetic susceptibility may lead to a faster accumulation of DNA damage and to a higher carcinogenic accumulation, resulting in an earlier onset of lung cancer. Biologic confirmation of this hypothesis is needed.

It has been demonstrated that polymorphisms in the genes involved in Phase 1 activation and Phase 2 detoxification of carcinogens alter lung cancer susceptibility.[6,22,28–30] Cumulative smoking and dietary factors were examined as potential confounders or as effect modifiers of the polymorphisms, whereby the association between pack-years or dietary habits and lung cancer differed among the different genotypes[6–8,16,22,28–37] Polymorphic variants of the cytochrome P450 enzymes (Phase 1) encoded by the *CYP* gene superfamily have been implicated in modifying lung cancer risk. The activated products of Phase 1 can bind to DNA (adducts), and among 1 of the more active carcinogens in tobacco smoke is the activated form of B(a)P, B(a)P-7,8-diole-9,10-epoxide (BPDE).

In contrast, Phase 2 enzymes detoxify activated carcinogens, rendering them hydrophilic and, thus, more excretable. The *GST* superfamily[38] has prominent Phase 2 function, and *GSTP1* is involved primarily in the detoxification of BPDE.[11–13] Studies have evaluated the effect of *GSTP1* on lung cancer risk. Ryberg and colleagues[12] reported a significantly higher frequency of the *GSTP1 GG* genotype among lung cancer patients than among controls (15.9% vs. 9.1%). The data from Ryberg et al., although not statistically significant, also suggest that there are higher levels of DNA adducts in patients who have the *GSTP1 GG* polymorphism compared with patients who have the *GSTP1 AA* polymorphism (15.5 ± 10.2 per 10^8 nucleotides vs. 7.9 ± 5.1 per 10^8 nucleotides).[14] Watson et al. reported differences in GST activity when comparing the *GSTP1 GG* genotype with *GSTP1 AA* genotype and reported specifically that enzymatic activity was significantly lower for the *GSTP1 GG* genotypes.[15]

The role of the heterozygote status often is unclear when examining genetic susceptibility to cancer. In the case of the *GSTP1* gene, there was no significant association between the heterozygote status (*AG*) and lung cancer. A review of the literature revealed no published statistically significant association between *GSTP1 AG* and lung cancer, although the actual estimate of the OR was < 1 in several studies.[7,14,15–19,39]

We evaluated the roles that disease stage and pack-years may play in the results. For disease stage, models with interaction between the genotype and age that compared the control group with the early-stage case group and separately with the late-stage case group were analyzed. The association of the genotype-age interaction and lung cancer risk did not differ. For pack-years, we ran stratified models to

determine whether the association of the genotype-age association and lung cancer risk also was dependent on cumulative smoking exposure. The strata were based on the median pack-years among the controls. The results were not different in heavy smokers or light smokers, as defined by pack-years.

We acknowledge several limitations to our current study. A concern of any case–control study is selection bias and control selection. In the current investigation, both patients and study investigators were blinded to genotypic status at the time of recruitment, and all genotyping was done blinded of case status. If there was a factor linked to the particular "at-risk" polymorphism that would lead to a higher directional recruitment of cases compared with controls, then there may be reason for concern about selection bias. This was not likely in the current study. Control selection is another concern. Our controls were recruited in the hospital but were not hospital patients themselves. Forty-one percent of controls were nonblood-related family and friends of the cases, whereas the remaining 59% of controls were family and friends of nonlung cancer patients at MGH (i.e., they had no relationship to any case). This design was adopted for its efficiency in recruiting controls. To address concerns of representativeness, we compared the smoking habits of our controls with the smoking habits of the general Massachusetts population and found no significance differences. We believe that these controls were likely to be referred to MGH for treatment if they ever became cases. Anecdotally, a very small number of controls have become cases seen at MGH—these individuals were eliminated from all analyses. Another limiting factor was the number of cases who were age <50 years at diagnosis ($n = 159$ patients), which was reduced further when they were stratified by genotype. However, all of our cases were recruited from the same study population.

We conclude that the *GSTP1 GG* genotype is associated with a higher risk of earlier onset lung cancer. Future investigators who study the role of genetic polymorphisms associated with a higher risk of cancer should consider the age of patients when performing their analysis. The effect of these polymorphisms may be more evident in younger populations. Although all smokers are at a much greater risk of lung cancer, the presence of the *GSTP1 GG* polymorphism may lead to an earlier presentation of disease.

REFERENCES

1. Kreuzer M, Kreienbrock L, Gerken M, et al. Risk factors for lung cancer in young adults. *Am J Epidemiol*. 1998;147: 1028–1037.
2. Kreuzer M, Kreienbrock L, Muller KM, Gerken M, Wichmann E. Histologic types of lung carcinoma and age at onset. *Cancer*. 1999;85:1958–1965.
3. Gadgeel SM, Ramalingam S, Cummings G, et al. Lung cancer in patients <50 years of age: the experience of an academic multidisciplinary program. *Chest*. 1999;115:1232–1236.
4. Cote ML, Kardia SL, Wenzlaff AS, Land SJ, Schwartz AG. Combinations of glutathione S-transferase genotypes and risk of early-onset lung cancer in Caucasians and African Americans: a population-based study. *Carcinogenesis*. 2005;26: 811–819.
5. Taioli E, Gaspari L, Benhamou S, et al. Polymorphisms in CYP1A1, GSTM1, GSTT1 and lung cancer below the age of 45 years. *Int J Epidemiol*. 2003;32:60–63.
6. Garcia-Closas M, Kelsey KT, Wiencke JK, Xu X, Wain JC, Christiani DC. A case-control study of cytochrome P450 1A1, glutathione S-transferase M1, cigarette smoking and lung cancer susceptibility (Massachusetts, United States) [published erratum appears in *Cancer Causes Control*. 1998;9: 126]. *Cancer Causes Control*. 1997;8:544–553.
7. Jourenkova-Mironova N, Wikman H, Bouchardy C, et al. Role of glutathione S-transferase GSTM1, GSTM3, GSTP1 and GSTT1 genotypes in modulating susceptibility to smoking-related lung cancer. *Pharmacogenetics*. 1998;8:495–502.
8. Le Marchand L, Sivaraman L, Pierce L, et al. Associations of CYP1A1, GSTM1, and CYP2E1 polymorphisms with lung cancer suggest cell type specificities to tobacco carcinogens. *Cancer Res*. 1998;58:4858–4863.
9. London SJ, Daly AK, Cooper J, Navidi WC, Carpenter CL, Idle JR. Polymorphism of glutathione S-transferase M1 and lung cancer risk among African-Americans and Caucasians in Los Angeles County, California. *J Natl Cancer Inst*. 1995; 87:1246–1253.
10. Saarikoski ST, Voho A, Reinikainen M, et al. Combined effect of polymorphic GST genes on individual susceptibility to lung cancer. *Int J Cancer*. 1998;77:516–521.
11. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*. 1999;91:1194–1210.
12. Fields WR, Morrow CS, Doss AJ, Sundberg K, Jernstrom B, Townsend AJ. Overexpression of stably transfected human glutathione S-transferase P1-1 protects against DNA damage by benzo[a]pyrene diol-epoxide in human T47D cells. *Mol Pharmacol*. 1998;54:298–304.
13. Nakajima T, Elovaara E, Anttila S, et al. Expression and polymorphism of glutathione S-transferase in human lungs: risk factors in smoking-related lung cancer. *Carcinogenesis*. 1995;16:707–711.
14. Ryberg D, Skaug V, Hewer A, et al. Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk. *Carcinogenesis*. 1997;18: 1285–1289.
15. Watson MA, Stewart RK, Smith GB, Massey TE, Bell DA. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. *Carcinogenesis*. 1998;19:275–280.
16. Kihara M, Noda K. Lung cancer risk of the GSTM1 null genotype is enhanced in the presence of the GSTP1 mutated genotype in male Japanese smokers. *Cancer Lett*. 1999;137: 53–60.
17. Harris MJ, Coggan M, Langton L, Wilson SR, Board PG. Polymorphism of the Pi class glutathione S-transferase in normal populations and cancer patients. *Pharmacogenetics*. 1998;8:27–31.

18. To-Figueras J, Gene M, Gomez-Catalan J, et al. Genetic polymorphism of glutathione S-transferase P1 gene and lung cancer risk. *Cancer Causes Control*. 1999;10:65–70.
19. Katoh T, Kaneko S, Takasawa S, et al. Human glutathione S-transferase P1 polymorphism and susceptibility to smoking related epithelial cancer; oral, lung, gastric, colorectal and urothelial cancer. *Pharmacogenetics*. 1999;9:165–169.
20. Massachusetts Department of Public Health. Massachusetts Tobacco Survey. Available at URL: http://www.state.ma.us/dph/mtcp/report/mats.htm (accessed May 20, 2006).
21. Wiencke JK, Thurston SW, Kelsey KT, et al. Early age at smoking initiation and tobacco carcinogen DNA damage in the lung [see comments]. *J Natl Cancer Inst*. 1999;91:614–619.
22. Xu X, Kelsey KT, Wiencke JK, Wain JC, Christiani DC. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev*. 1996;5:687–692.
23. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. New York: John Wiley & Sons, Inc.; 2000.
24. Bartsch H, Nair U, Risch A, Rojas M, Wikman H, Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*. 2000;9:3–28.
25. SAS Institute, Inc. SAS/Stat User's Guide. Version 6, 4th ed. Cary, NC: SAS Institute, Inc.; 1998.
26. Bailey-Wilson JE, Sellers TA, Elston RC, Evens CC, Rothschild H. Evidence for a major gene effect in early-onset lung cancer. *J La State Med Soc*. 1993;145:157–162.
27. Gauderman WJ, Morrison JL. Evidence for age-specific genetic relative risks in lung cancer. *Am J Epidemiol*. 2000;151:41–49.
28. Yang P, Schwartz AG, McAllister AE, Swanson GM, Aston CE. Lung cancer risk in families of nonsmoking probands: heterogeneity by age at diagnosis. *Genet Epidemiol*. 1999;17:253–73.
29. Hengstler JG, Arand M, Herrero ME, Oesch F. Polymorphisms of N-acetyltransferases, glutathione S-transferases, microsomal epoxide hydrolase and sulfotransferases: influence on cancer susceptibility. *Recent Results Cancer Res*. 1998;154:47–85.
30. Taioli E, Ford J, Trachman J, Li Y, Demopoulos R, Garte S. Lung cancer risk and CYP1A1 genotype in African Americans. *Carcinogenesis*. 1998;19:813–817.
31. Benhamou S, Reinikainen M, Bouchardy C, Dayer P, Hirvonen A. Association between lung cancer and microsomal epoxide hydrolase genotypes. *Cancer Res*. 1998;58:5291–5293.
32. Cheng TJ, Christiani DC, Xu X, Wain JC, Wiencke JK, Kelsey KT. Glutathione S-transferase mu genotype, diet, and smoking as determinants of sister chromatid exchange frequency in lymphocytes. *Cancer Epidemiol Biomarkers Prev*. 1995;4:535–542.
33. Le Marchand L, Murphy SP, Hankin JH, Wilkens LR, Kolonel LN. Intake of flavonoids and lung cancer. *J Natl Cancer Inst*. 2000;92:154–160.
34. London SJ, Daly AK, Fairbrother KS, et al. Lung cancer risk in African-Americans in relation to a race-specific CYP1A1 polymorphism [see comments]. *Cancer Res*. 1995;55:6035–6037.
35. Nyberg F, Hou SM, Hemminki K, Lambert B, Pershagen G. Glutathione S-transferase mu1 and N-acetyltransferase 2 genetic polymorphisms and exposure to tobacco smoke in nonsmoking and smoking lung cancer patients and population controls. *Cancer Epidemiol Biomarkers Prev*. 1998;7:875–883.
36. Persson I, Johansson I, Lou YC, et al. Genetic polymorphism of xenobiotic metabolizing enzymes among Chinese lung cancer patients. *Int J Cancer*. 1999;81:325–329.
37. To-Figueras J, Gene M, Gomez-Catalan J, et al. Glutathione-S-transferase M1 and codon 72 p53 polymorphisms in a northwestern Mediterranean population and their relation to lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev*. 1996;5:337–342.
38. Casarett LJ, Klaassen CD, Amdur MO, Doull J. Casarett and Doull's Toxicology: the Basic Science of Poisons. New York: McGraw-Hill Health Professions Division; 1996.
39. Harries LW, Stubbins MJ, Forman D, Howard GC, Wolf CR. Identification of genetic polymorphisms at the glutathione S-transferase Pi locus and association with susceptibility to bladder, testicular and prostate cancer. *Carcinogenesis*. 1997;18:641–644.