The *p53* Codon 72 Polymorphism and Lung Cancer Risk[1]

Rong Fan, Ming-Tsang Wu, David Miller, John C. Wain, Karl T. Kelsey, John K. Wiencke, and David C. Christiani[2]

Departments of Environmental Health (Occupational Health Program) [R. F., M-T. W., D. M., K. T. K., D. C. C.] and Cancer Cell Biology [K. T. K.], Harvard School of Public Health, Boston, Massachusetts 02115; Thoracic Surgery Unit, Department of Surgery [J. C. W.], and Pulmonary and Critical Care Unit, Department of Medicine [D. C. C.], Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts 02114; University of California at San Francisco, Department of Epidemiology and Biostatistics, San Francisco, California 94143 [J. K. W.]

Abstract

The *p53* tumor suppressor gene frequently is mutated in many forms of human carcinomas. A common polymorphism occurs at codon 72 of exon 4, with two alleles encoding either arginine (CGC) or proline (CCC). This *p53* polymorphism reportedly is associated with lung cancer susceptibility. However, not all investigations have been consistent, and this hypothesized association remains controversial. We tested the hypothesis that the *Pro/Pro* genotype is associated with increased lung cancer risk in a large case-control study of lung cancer that included 482 cases and 510 controls from the Massachusetts General Hospital in Boston, Massachusetts. DNA from peripheral blood samples was examined by PCR-RFLP. *Pro/Pro* homozygotes were found more frequently in adenocarcinomas (cases, 16.4%; controls, 12.0%; $P = 0.03$). The prevalence of the *Pro/Pro* homozygous genotype increased in frequency with increasing pack-years of smoking. The combined susceptible genotype homozygous *Pro/Pro* and heterozygous *Arg/Pro* was associated with a 1.45-fold higher risk of adenocarcinoma compared with *Arg/Arg* genotype (95% confidence interval = 1.01–2.06; $P = 0.04$) after adjustment for relevant variables. Lung adenocarcinoma risk increased with the presence of one or both variant alleles across smoking strata. In addition, at each level of smoking (except nonsmoker and light smoker), the risk associated with smoking was higher for the population with the combined variant (*Arg/Pro* + *Pro/Pro*) genotype. The risk for the combined genotype was associated with tobacco exposure status. In conclusion, the codon 72 germ-line polymorphism (*Arg/Pro*) of the common tumor suppressor gene *p53* contributes to heritable susceptibility for smoke-induced lung adenocarcinoma. The modifications by *p53* polymorphism and pack-years resulted in an increased risk of the susceptible genotype to lung adenocarcinoma. The *p53* gene may modulate the response to environment carcinogens and thereby affect the risk of developing lung adenocarcinoma.

Introduction

The *p53* tumor suppressor gene, located on chromosome 17p13, is one of the most commonly mutated genes in all types of human cancer (1, 2). Recent studies of the function of the wild-type *p53* demonstrated that its antiproliferative effect is mediated by stimulation of a 21-kDa protein (p21cip1/waf1) that inhibits cyclin-dependent kinase activity and, thereby, cell division (3, 4). This negative cell cycle controller effect may explain why the wild-type *p53* gene can suppress the transformation of cells by activated oncogenes, thereby inhibiting the growth of malignant cells *in vitro* and suppressing the tumorigenic phenotype *in vivo* (5, 6). Analysis of somatic tissue from many human cancers has shown that the wild-type *p53* allele frequently is lost and a mutant allele retained, providing a growth advantage for malignant cells (7–9). The mutation of the *p53* gene can damage its DNA-binding properties and transcription factor function, inhibiting its normal function in cell cycle control and in cell proliferation (10).

To date, several polymorphisms in the wild-type *p53* gene locus have been described. The codon 72 polymorphism on the 4th exon of the *p53* gene, which produces variant proteins with an arginine (CGC) or proline (CCC), has been reported to be associated with bladder and lung cancer (11–13).

An association of the codon 72 *p53* polymorphism with lung cancer susceptibility has been reported by several authors. In one study, Weston *et al.* (13) reported an increased frequency of the proline allele in adenocarcinomas, but a later study by this same group (14) did not confirm this finding in a different set of cancer cases and controls. The homozygous *Pro/Pro* genotype was found to be overrepresented in a study of Japanese lung cancer, especially in Kreyberg type I but not in adenocarcinoma (15). More recently, an enhanced risk was reported for African Americans with both the *Pro/Pro* genotype and an early onset of lung cancer (16). A Swedish study has also suggested that the codon 72 alleles may not be functionally involved in lung cancer but, rather, may be a marker in linkage dysequilibrium with other cancer susceptibility sites (17). A Spanish study reported that there was no difference in the prevalence of the codon 72 *p53* polymorphism between lung cancer cases and controls. However, in that study, the *Pro* allele of the *p53* germ-line polymorphism increased slightly the lung cancer risk for the *GSTM1*-null genotype among smokers (18). Murata *et al.* (19) reported that, among 191 lung cancer cases, 115 colorectal cancer patients, and 152 controls, there was a statistically significant difference in genotype frequency only in nonsmokers with lung cancer, with the homozygous *Arg/Arg* genotype overrepresented in that group.

Hence, the literature to date has not been consistent with respect to the association of codon 72 polymorphism with lung

Received 1/11/00; revised 6/14/00; accepted 8/7/00.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Supported by NIH Grants CA74386, ES/CA06409, ES8357, and ES00002.
[2] To whom requests for reprints should be addressed, at Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115. E-mail: dchris@hohp.harvard.edu.

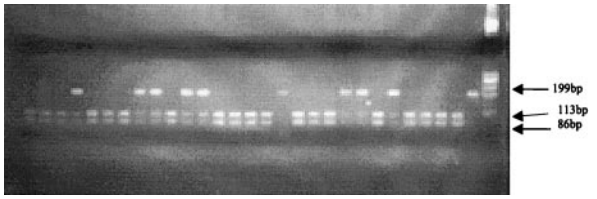

Fig. 1. PCR-RFLP analysis of the *p53* gene. The *Pro* allele is not cleaved by *Bst*UI at codon 72 and has a single band with a fragment length of 199 bp. The *Arg* allele is cleaved by *Bst*UI and yields two small fragments (113 and 86 bp). The heterozygote has three bands (199, 113, and 86 bp).

cancer susceptibility. In the present study, we conducted a large case-control study of lung cancer patients and controls and examined the genotype frequency of codon 72 of lung cancer patients and controls, using PCR-based genotyping methods to further evaluate the possible relevance of this polymorphism for lung cancer risk.

Materials and Methods

Subjects. The present study was a hospital-based case-control study that included 482 lung cancer patients and 510 controls. Eligible cases included all patients with primary lung cancer (stages I and II) presenting for thoracic surgery at the Massachusetts General Hospital between December 1992 and August 1996. Controls ($n = 510$) were friends or spouses of the following groups: lung cancer ($n = 190$), other cardiothoracic surgery patients ($n = 320$).

A detailed interviewer-administered questionnaire was completed for each case and control by a trained interviewer. A modified standardized American Thoracic Society respiratory questionnaire (20) with additions on a detailed occupational and environment exposure history was used. The questionnaire included information on average cigarettes smoked daily for current smokers, years smoked, and time since quitting smoking for ex-smokers. As an indication of cumulative smoking exposure, pack-years were defined as the average number of packs smoked per day multiplied by years smoked. We also obtained information on all of an individual's job titles, job tasks, years and dates of exposure, and family history of cancer in first-degree relatives.

***p53 Bst*UI Polymorphism.** PCR-RFLP analysis of the codon 72 of the *p53* gene originally described by Ara *et al.* (1) was used to identify *p53 Bst*UI genotypes. The two primers were 5′-TTGCCGTCCCAAGCAATGGATGA-3′ and 5′-TCTGGGAAGGGACAGAAGATGAC-3′. Each PCR reaction mixture (50 μl) contained 10 pmol of each primer, 2.0 mM $MgCl_2$, 200 mM each dNTP, 1 unit of *Taq* polymerase and 100–300 ng of genomic DNA. Reaction mixtures were preincubated for 5 min at 94°C. PCR conditions were 94C° for 30 s and 55°C for 1 min, followed by 72°C for 1 min for 35 rounds. After confirmation of an amplified fragment of the expected size (199 bp) on an agarose gel, the PCR products were digested with 2 units of restriction enzyme *Bst*UI (New England Biolabs, Beverly, MA) at 60°C for 16 h. DNA fragments were electrophoresed through a 2% agarose gel and stained with ethidium bromide (Fig. 1).

Statistical Analysis. Univariate statistics (χ^2 and *t* tests) were used first to compare cases and controls for demographic variables and genotype prevalence. Multivariate logistic regression analysis was performed to assess the association between the *p53* polymorphism and lung cancer. Potential confounding factors adjusted for included sex, age (years), race, education level, smoking status (current, ex-smoker, nonsmoker) and pack-years. For the purpose of modeling the association between the *p53* variant gene frequency and lung cancer, we compared the homozygous (*Pro/Pro*) variant and the heterozygous genotype (*Arg/Pro*) with the homozygous (*Arg/Arg*) genotype, respectively. Because of low number of homozygous *Pro/Pro* subjects in some subgroups based on pack-years, we combined the homozygous *Pro/Pro* variant with heterozygous *Arg/Pro* as a single group to compare with the homozygous *Arg/Arg* genotype. In this model, we assumed that risk associated with the *Arg/Pro* genotype would be intermediate between that of the *Arg/Arg* and the *Pro/Pro* genotypes. On the basis of this assumption, we coded the data as follows (using logistic regression model): *Arg/Arg* genotype = 0; *Arg/Pro* genotype = 0.5; *Pro/Pro* genotype = 1. The resulting coefficient yields the risk associated with having the *Pro/Pro* + *Arg/Pro versus Arg/Arg* genotype, adjusting for covariates.

Table 1 Distribution of selected variables between lung cancer cases and controls

Characteristic	Cases ($n = 482$)	Controls ($n = 510$)
Sex, *n* (%)		
M	264 (54.9)	229 (45.1)
F	217 (45.1)	279 (54.9)
Age, mean (SD), years	65.46 (10.2)	61.6 (10.9)
Smoking status, *n* (%)		
Never[a]	22 (4.6)	155 (30.5)
Ex-smoker[b]	257 (53.5)	262 (51.6)
Current smoker[c]	201 (41.9)	91 (17.9)
Pack-years, mean (SD)	57.77 (39.3)	22.62 (27.3)
Cigarettes per day, mean (SD)[d]	28.64 (17.0)	15.34 (15.4)
Education, *n* (%)		
College graduate	100 (21.5)	112 (22.4)
<College graduate	366 (78.5)	387 (77.6)
Race, *n* (%)		
Caucasian	466 (95.7)	497 (97.8)

[a] Less than 0.05 pack-years in lifetime.
[b] Quit smoking at least 1 year before enrollment.
[c] Smoked at least one cigarette per day for at least 1 year or 20 packs of cigarettes or 12 ounces of tobacco in rolled cigarettes in lifetime.
[d] Cigarettes per day among smokers.

Results

The distribution of demographic variables for cases and controls is summarized in Table 1. Men are overrepresented in the cases (54.9 *versus* 45.1%), with females overrepresented in the controls (54.9 *versus* 45.1%). The mean ages were 65.5 years for cases and 61.6 years for controls. Predictably, cases were significantly more likely to be current smokers (41.9 *versus* 17.9%), to have smoked more cigarettes per day (28.6 *versus* 15.4), and to have accumulated more pack-years (57.8 *versus* 22.6) than the controls. There was no statistically significant difference between cases and controls with regard to education and race.

We examined of the distribution of *p53* codon 72 genotypes among controls and lung cancer patients by histological subtypes (Table 2). The frequencies of the three genotypes, *Arg/Arg*, *Arg/Pro*, and *Pro/Pro*, were 46.5, 41.6, and 12.0%, respectively, in controls. The crude genotypic frequencies in the lung cancer patients were similar to those of the controls. When lung cancer cases were stratified by histological subtype, the distribution of the three genotypes in adenocarcinoma patients

Table 2 Frequency of *p53* genotypes among controls and histological types of lung cancer

	n	*Arg/Arg*, *n* (%)	*Arg/Pro*, *n* (%)	*Pro/Pro*, *n* (%)
Controls	510	237 (46.47)	212 (41.57)	61 (11.96)
Hystological types				
All	482	212 (43.98)	204 (42.32)	66 (13.69)
Adenocarcinoma	244	96 (39.34)	108 (44.26)	40 (16.39)[a]
Squamous cell	133	73 (54.89)	44 (33.08)	16 (12.03)
Large cell	26	11 (42.31)	13 (50)	2 (7.69)
Small cell	25	9 (36)	13 (52)	3 (12)
Broncho-alveolar	10	5 (50)	4 (40)	1 (1)
Mixed	16	7 (43.75)	8 (50)	1 (6.25)
>1 primary lung cancer	21	6 (28.57)	12 (57.14)	3 (14.29)

[a] $P = 0.03$ for comparison of *p53 Pro/Pro* genotype with *p53 Arg/Arg* genotype between adenocarcinoma cases and controls.

differed from controls: *Arg/Arg*, *Arg/Pro*, and *Pro/Pro* were 39.3, 44.3, and 16.4%, respectively. There was no statistically significant difference in the prevalence of the polymorphism among squamous cell, large cell, small cell, mixed cell, and bronchoalveolar cell subtypes.

We examined the data further by stratifying values by potentially important confounding factors (Table 3). Within lung adenocarcinoma patients and controls, the variant *Pro/Pro* genotype was strongly associated with cigarette smoking, with a higher prevalence in current smokers. Among cases, but not controls who smoked, the prevalence of the variant genotype appeared to vary by cumulative cigarette consumption category, with a lower prevalence of the *Pro/Pro* genotype in light smokers but a higher prevalence in moderate and heavy smokers among adenocarcinoma cases but not controls. In older cases (65–71 years), the prevalence of the *Pro/Pro* genotype was higher than in controls (not statistically significant). The other variables examined were not statistically significant between cases and controls.

We then examined the association of *p53* genotype with lung adenocarcinoma, using a multivariate logistic regression model (Table 4). We first compared the *Pro/Pro* genotype with the *Arg/Arg* genotype. The crude OR[3] for *Pro/Pro versus Arg/Arg* was 1.61 ($P = 0.03$; 95% CI, 1.04–2.49). After adjustment for age, sex, race, education, pack-years, and smoking status, the same covariates, the OR remained elevated 1.59 ($P = 0.06$; 95% CI, 0.97–2.61). For the purpose of this analysis, the *Pro/Pro* genotype was combined with the *Arg/Pro* genotype because of low prevalence of the *Pro/Pro* in subgroups of subjects based on pack-years. The combined variant genotype group had a 1.34-fold higher risk for adenocarcinoma than the genotype *Arg/Arg* ($P = 0.06$; 95% CI, 0.98–1.83). After adjustment for the same covariates, the OR was 1.45 ($P = 0.04$; 95% CI, 1.02–2.06). Because we assumed the genetic contribution to risk might be less at very high doses of smoking (21), we excluded the group that included the top 15% of subjects based on pack-years (84 pack-years) and repeated the same logistic regression model (data not shown). The OR (both crude and adjusted) rose slightly, with an adjusted OR for the combined variant genotype of 1.50 (95% CI = 1.03–2.18) and an adjusted OR for the *Pro/Pro* genotype of 1.79 (95% CI = 1.06–3.02). To avoid the effect of ethnicity, we ran the logistic model two ways (with all races or with only Caucasians), but the results remained almost the same.

Table 3 Frequency (%) of the *p53* variant among lung adenocarcinoma cases and controls

Variable	Adenocarcinoma		Controls		P^a
	n	*Pro/Pro*, *n* (%)	*n*	*Pro/Pro*, *n* (%)	
Sex[b]					
M	117	23 (19.7)	229	30 (13.1)	
F	126	17 (13.5)	279	31 (11.1)	0.41
Race[c]					
Caucasian	232	37 (16.0)	497	60 (12.1)	
Age (years)[d]					
<54	49	6 (12.2)	139	14 (10.1)	
54–64	60	8 (13.3)	127	14 (11.0)	
65–71	58	13 (22.4)	130	17 (13.1)	
>71	76	13 (17.1)	112	16 (14.3)	0.7
Education[e]					
College	59	9 (15.3)	112	17 (15.2)	
<College	177	29 (16.4)	387	43 (11.1)	0.61
Smoking status[f]					
Never	17	2 (11.8)	155	15 (9.7)	
Ex-smoker	129	21 (16.3)	262	33 (12.6)	
Current	97	17 (17.5)	91	13 (14.3)	0.01
Pack-years[g]					
0	17	2 (11.8)	155	15 (9.7)	
1–29	57	5 (8.8)	184	24 (13.0)	
29–56	84	18 (21.4)	114	14 (12.3)	
>56	85	15 (17.7)	54	7 (13.0)	<0.0001

[a] For trends comparing adenocarcinoma with controls.
[b] Information was missing for 1 case and 2 controls.
[c] Information was missing for 1 case and 2 controls.
[d] Information was missing for 1 case and 2 controls.
[e] Information was missing for 8 cases and 11 controls.
[f] Information was missing for 1 case and 2 controls.
[g] Information was missing for 1 case and 3 controls.

The data in Table 4 display the association of the combined contributions of genotype and smoking exposure to lung adenocarcinoma risk, where the *Arg/Arg* genotype and no pack-years (nonsmokers) was used as the referent group. Pack-years were divided into quartiles, and the *Pro/Pro* and *Arg/Pro* genotypes (examined) were one category. As expected, lung adenocarcinoma risk increased with increasing cumulative cigarette dose. In addition, at each level of smoking except the light smokers, the risk associated with smoking was higher for persons with the *Pro/Pro* + *Arg/Pro* combined genotype. The group with both the heaviest tobacco smoke exposure and the *p53* combined *Pro/Pro* + *Arg/Pro* genotype had the highest risk of lung adenocarcinoma, roughly 39-fold higher (95% CI, 10.76–140.51) than the lowest risk group of nonsmokers who were *p53 Arg/Arg*. The results revealed the same trend when quintiles or sextiles of pack-years were used. Because the number of cases who had never smoked was too low ($n = 17$), we combined those who had never smoked with subjects who smoked 1–29 pack-years into a single group and reran the logistic regression model; the results showed the same trend. As expected, the adjusted OR for lung adenocarcinoma by *p53* and pack-years decreased, and the 95% CI became tighter (Table 5).

Discussion

There is an expanding body of literature suggesting that host factors, including genetic polymorphisms, may explain some of the individual differences in cancer occurrence (22, 23). We have shown that the codon 72 germ-line polymorphism (*Arg-*

[3] The abbreviations used are: OR, odds ratio; CI, confidence interval.

Table 4 Adjusted ORs for lung adenocarcinoma by *p53* genotype and pack-years

Nonsmokers with the *p53 Arg/Arg* genotype were used as the reference group to show the adjusted OR of lung adenocarcinoma for a given pack-year and within *p53* (*Arg/Arg*, *Arg/Pro* + *Pro/Pro/Pro*) genotype stratas.

		p53 genotype	
		Arg/Arg	*Arg/Pro* + *Pro/Pro*
Total	Cases	96	148
	Controls	237	273
	OR (95% CI)[a]	1 (Ref.)[b]	1.34 (0.98–1.83)
	OR (95% CI)[c]	1 (Ref.)	1.45 (1.02–2.06)
Stratum, based on pack-years			
0	Cases	5	12
	Controls	72	83
	OR[d]	1	3.1
	95% CI	Ref.	0.8–11.5
1–29	Cases	25	32
	Controls	78	106
	OR[d]	7.3	7.2
	95% CI	2.1–25.8	2.1–24.9
30–56	Cases	31	53
	Controls	60	54
	OR[c]	10.9	21.8
	95% CI	3.1–38.8	6.2–76.2
>56	Cases	35	50
	Controls	25	29
	OR[d]	34.1	38.9
	95% CI	9.2–126.1	10.7–140.5

[a] Crude OR.
[b] Ref., within reference values.
[c] Adjusted for sex, age, education, smoking status, pack-years.
[d] Adjusted for sex, age, education, smoking status.

Table 5 Adjusted ORs for lung adenocarcinoma by *p53* genotype and pack-years

Nonsmokers with the *p53 Arg/Arg* genotype were used as the reference group to show the adjusted OR for lung adenocarcinoma for a given pack-year and within *p53* (*Arg/Arg*, *Arg/Pro* + *Pro/Pro*) genotype stratas.

		p53 genotype	
		Arg/Arg	*Arg/Pro* + *Pro/Pro*
Total	Cases	96	148
	Controls	237	273
	OR (95% CI)[a]	1 (Ref.)[b]	1.34 (0.98–1.83)
	OR (95% CI)[c]	1 (Ref.)	1.45 (1.02–2.06)
Stratum, based on pack-years			
0–29	Cases	30	44
	Controls	150	189
	OR[d]	1	1.2
	95% CI	Ref.	0.7–2.2
30–56	Cases	31	53
	Controls	60	54
	OR[d]	1.7	3.6
	95% CI	0.9–3.4	1.9–6.8
>56	Cases	35	50
	Controls	25	29
	OR[d]	5.4	6.1
	95% CI	2.6–11.3	3.1–12.2

[a] Crude OR.
[b] Ref., within reference values.
[c] Adjusted for sex, age, education, smoking status, pack-years.
[d] Adjusted for sex, age, education, smoking status.
0–29 includes nonsmoker and pack-year <29 smokers.

Pro polymorphism) of the common tumor suppressor *p53* gene contributes to susceptibility to smoking-induced adenocarcinoma of the lung. The *Pro/Pro* homozygous genotype occurred more frequently in adenocarcinomas. The prevalence of the *Pro/Pro* genotype in adenocarcinoma was higher than that of other genotypes and increased with increasing pack-years.

In this study, we examined the prevalence of *p53* codon 72 polymorphisms in a Caucasian group of lung cancer patients and controls. The prevalence of the *Pro/Pro* genotype in adenocarcinoma cases was statistically different from that of the controls (16.4 *versus* 12.0%). A Japanese study reported that the prevalence of the *Pro/Pro* variant in patients with adenocarcinoma was 1.2-fold higher, which was not statistically different from controls (15). Weston *et al.* (13) reported a prevalence of 26% for their pooled control group and 21% in all types of lung cancer combined. Jin *et al.* (16) reported that the susceptible *Pro/Pro* genotype was associated with a 1.6-fold higher risk of all types of lung cancer combined in African Americans and 1.9-fold higher risk in Mexican Americans, with neither reaching statistical significance. Weston *et al.* (14) later reported no association between the allele frequencies of *p53* and susceptibility for all lung cancers in a Caucasian and an African-American population. The discordance in these studies may be the result of choice of controls, the small sample sizes, and hence, the inability to stratify by histological type.

In our study, the frequency of the *Pro/Pro* genotype in adenocarcinoma was much lower than that of the *Arg/Arg* genotype in the low pack-years stratum. The prevalence of the *Pro/Pro* genotype rose linearly with pack-years. Our results show that the frequency of the *Pro/Pro* genotype was low at lower cumulative cigarette levels, in contrast to several other published reports (13–15). However, the major reason for the low prevalence of the *Pro/Pro* genotype in our study was that, at light cumulative levels, the number of subjects was small. Because the homozygous *Pro/Pro* variant frequency is so low in some pack-years quintiles, we reanalyzed the data after combining the heterozygous and homozygous variants and found a statistically significant association with adenocarcinoma. Furthermore, we found that the most important confounding factor for the study appears to be pack-years, which must be adequately adjusted in the analysis of association between *p53* genotype and lung adenocarcinoma. The association of the *p53* codon 72 variant genotype with increased lung adenocarcinoma risk was modified by gender, race, age, education, smoking status, and pack-years and remained statistically significant after we adjusted for these factors.

As expected, the OR for lung adenocarcinoma increased with dose of cigarette tobacco smoke for both genotypes. In addition, at each level of smoking except for light smokers, subjects with the *p53* combined variant genotype experienced higher risk of adenocarcinoma than subjects who had the *p53 Arg/Arg* genotype. Individuals with the heaviest tobacco smoke exposure and *p53* combined genotype had the highest risk of lung cancer, roughly 38-fold higher than the lowest risk group of nonsmokers who were *Arg/Arg*. These data suggest that the presence of the *p53* gene product exerts a protective effect for smoking-induced lung cancer and that the presence of one variant allele alters this product. Thus, the modifications by *p53* and pack-years work independently and increase the risk of the susceptible genotypes for lung adenocarcinoma. Considering the biological role of *p53* in carcinogenesis, one of the most plausible interpretations for the distribution of the *Pro/Pro* genotype is that this heritable polymorphism imparts high risk of developing adenocarcinoma.

Genetic differences in risk may be smaller at high loads of carcinogen exposure, when environmental influences may overcome the association with a genetic predisposition (21, 24,

25). However, in our study, we found that there was little difference between the susceptible genotypes and pack-years at higher doses of smoking.

The reason for the observed tissue-specific difference in the risk conferred by the germ-line *p53* polymorphism is unknown. There have been changing trends in the occurrence of adenocarcinoma of the lung over the past two decades. It has been hypothesized that changes in cigarette composition and smoking behavior have produced a histological shift over the past decades. The proportional decrease in lung small cell carcinomas may be related to a reduction in exposure of the central bronchi to polyaromatic hydrocarbons, and the dramatic rise in lung adenocarcinoma may be related to increased exposure of the peripheral lung to tobacco-specific nitrosamines. The hypothesis suggested by Kawajiri *et al.* (23) is that different carcinogenic processes are involved in the genesis of various tumor types because of the presence of functionally different *p53* alleles (*Pro*- or *Arg*-type). The functional difference of the *p53* polymorphism at codon 72 has been reported (26). A *p53 Arg/Arg* genotype induces apoptosis with faster kinetics and suppresses transformation more efficiently than the *p53 Pro/Pro* genotype.

Our observations provide evidence that cigarette smoking should be appropriately controlled in the analysis of *p53* codon 72 polymorphisms and lung cancer and that when adjusted for smoking, the presence of the codon 72 variant allele is associated with increased lung adenocarcinoma risk. The previously reported inconsistent findings for an association of the codon 72 variant in the *p53* gene with lung cancer between light and heavy smokers could be attributable to confounding factors, to sample size limitations, and to various choices of controls.

Our study has several potential limitations that might have influenced our results. First, our controls come from friends or spouses referred by lung cancer cases or by cardiothoracic patients. A necessary condition for the validity of our OR estimates of the genotype association is that, conditional on the variables controlled in the analysis, (*a*) the genotype distribution of the two kinds of controls is the same and (*b*) represents the distribution in the source population. To evaluate whether the first statement is satisfied, we tested for differences in the genotype frequencies of the two control groups, finding no significant differences ($\chi^2 = 0.4$; $P = 0.5$). Given that we found no significant differences in genotype frequencies and that the frequency in the collapsed control group was similar to the estimates observed in past studies of Caucasian populations, it seems reasonable to assume that the two control groups represent a common population. Selecting controls among friends or spouses of lung cancer patients or other surgical patients might have introduced bias by two different mechanisms (27). First, the distribution of exposures among gregarious subjects, *i.e.*, subjects who tend to be named by more than one other person as controls, might be different from the distribution among non-gregarious subjects and, therefore, it might not represent the distribution in the source population that includes both types of subjects. Second, exposure among friends and spouses of patients might be similar to the exposure among the patients themselves, and thus controls may not represent the distribution in the source population. These two factors are likely to be relevant for variables such as age, gender, smoking, or dietary habit. However, within categories of these variables and among the same racial or ethnic group, the genetic polymorphisms in *p53* are unlikely to be associated with being gregarious or with having a friend or spouse with lung cancer. Therefore, the genotype-disease association is unlikely to be affected substantially by this type of bias.

On the other hand, the reported estimates for the effect of pack-years on lung cancer are likely to be underestimated. Moreover, our case population includes only surgically treated lung cancer patients (stages I and II), and those with more advanced tumors not eligible for surgery were not included. Thus, differential selection of cases with respect to the genotype could occur if the genotype is related to the stage at which the lung cancer is detected. To our knowledge, one Chinese study reported that patients with the *Pro/Pro* genotype were more than five times more likely to die at early postoperative stages than those with the *Arg/Pro* genotype (28). This may introduce overestimation of the genotype OR because poorly differentiated tumors and tumor with metastasis are more likely to become inoperable. Finally, ethnicity could be a confounder of the OR for the genotype-disease association. However, in the context of our study, this type of confounding is unlikely to be important because we got almost the same results when we included all races or just Caucasian subjects in the logistic regression model. Therefore, ethnicity is unlikely to be strongly associated either with the genotype, as indicated above, or with lung adenocarcinoma risk.

Our analysis of the possible association between the *p53* genotype and pack-years revealed that the susceptible genotype is more clearly associated with an increased risk of adenocarcinoma among smokers. This finding is in agreement with some reports (29, 30) and is discordant with others (15), and suggests that genetic susceptibility may play a role in certain medium exposure conditions but may be overpowered by heavy exposure to the carcinogens in high doses of tobacco smoke. Seemingly moderate genetic risk, combined with environmental exposures, can determine an important proportion of lung cancers. However, further study in defined light smoking subgroups is required to substantiate the findings that the *Pro/Pro* genotype predisposes individuals to adenocarcinoma of the lung.

Acknowledgments

We gratefully acknowledge the assistance of Linda Lineback, Barbara Bean, Li Su, Lilian Xu, Zheng-fa Zuo, and Lucy-Ann Principe-Hasan.

References

1. Ara, S., Lee, P. S., Hansen, M. F., and Saya, H. Codon 72 polymorphism of the Tp53 gene. Nucleic Acids Res., *18:* 4961, 1990.

2. Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. *p53* mutations in human cancers. Science (Washington DC), *253:* 49–53, 1991.

3. Marx, J., Momand, J., and Finlay, C. A. How *p53* suppresses cell growth. Science (Washington DC), *262:* 1644–1645, 1993.

4. Dulic, V., Kaufmann, W. K., Wilson, S. J., and Reed, S. I. *p53* dependent inhibition of cyclin-dependent kinase activities in human fibroblast during radiation-induced G_1 arrest. Cell, *76:* 1013–1023, 1994.

5. Weinberg, R. A. Tumor suppressor genes. Science (Washington DC), *254:* 1138–1146, 1991.

6. Frebourg, T., and Friend, S. H. The importance of *p53* gene alterations in human cancer: is there more than circumstantial evidence? J. Natl. Cancer Inst., *85:* 1554–1557, 1993.

7. Vogelstein, B., and Kinzler, K. W. *p53* function and dysfunction. Cell, *70:* 523–526, 1993.

8. Jensen, R., A., and Page, D. L. *p53*: the promising story continues to unfold. Hum. Pathol., *24:* 455–456, 1993.

9. Harris, C. C. *p53*: at the crossroads of molecular carcinogenesis and risk assessment. Science (Washington DC), *262:* 1980–1981, 1993.

10. Roy, B., Beamen, J., Balint, E., and Reisman, D. Transactivation of the human *p53* tumor suppressor gene by c-myc/max contributes to elevated mutant *p53* expression in some tumors. Mol. Cell. Biol., *14:* 7805–7815, 1994.

11. Oka, K., Ishkawa, J., Bruner, J. M., Takahashi, R., and Saya, H. Detection of loss of heterozygosity in the *p53* gene in renal cell carcinoma and bladder cancer using the polymerase chain reaction. Mol. Carcinog., *4:* 10–13, 1991.

12. Matlashewski, G. J., Tuck, S., Pim, D., Lamb, P., Schneider, J., and Crawford, L. V. Primary structure polymorphism at amino acid residue 72 of human *p53*. Mol. Cell. Biol., *7:* 961–963, 1987.

13. Weston, A., Perrin, L. S., Forrester, K., Hoover, R. N., Trump, B. F., Harris, C. C., and Caporaso, N. E. Allelic frequency of a *p53* polymorphism in human lung cancer. Cancer Epidemiol. Biomark. Prev., *1:* 481–483, 1992.

14. Weston, A., Ling-Cawley, H. M., Caporaso, N. E., Bowman, E. D., Hoover, R. N., Trump, B. F., and Harris, C. C. Determination of the allele frequencies of an L-myc and a *p53* polymorphism in human lung cancer. Carcinogenesis (Lond.), *15:* 583–587, 1994.

15. Kawajiri, K., Nakachi, K., Imai, K., Watanabe, J., and Hayashi, S. I. Germ line polymorphism of *p53* and *CYP1A1* genes involved in human lung cancer. Carcinogenesis (Lond.), *14:* 1085–1089, 1993.

16. Jin, X., Wu, X., Roth, J. A., Amos, C. I., King, T. M., Branch, C., Honn, S. E., and Spitz, M. R. Higher lung cancer risk for younger African-Americans with the *Pro/Pro p53* genotype. Carcinogenesis (Lond.), *16:* 2205–2208, 1995.

17. Birgander, R., Sjalander, A., Rannug, A., Alexandire, A. K., Ingelmansundberg, M., Seidegard, J., Tornling, G., Beckman, G., and Beckman, L. *P53* polymorphisms and haplotypes in lung cancer. Carcinogenesis (Lond.), *16:* 2233–2236, 1995.

18. To-Figueras, J., Gene, M., Gomez-Catalan, J. G., Galan, C., Firvida, J., Fuentes, M., Rodamilans, M., Huguet, E., Estape, J., and Corbella, J. Glutathione-*S*-Transferase M1, and codon 72 *p53* polymorphism in a northwestern Mediterranean population and their relation to lung cancer susceptibility. Cancer Epidemiol. Biomark. Prev., *5:* 337–342, 1996.

19. Murata, M., Tagawa, M., Kimura, M., Kimura, H., Watunabe, S., and Saisho, H. Analysis of a germ line polymorphism of the *p53* gene in lung cancer patients; discrete results with smoking history. Carcinogenesis (Lond.), *17:* 261–264, 1996.

20. Ferris, B. G. Epidemiology standardized project. Am. Rev. Respir. Dis., *118* (Suppl.)*:* 1s–120s, 1978.

21. Amos, C., Caporaso, B., and Weston, A. Host factors in lung cancer risk: a review of interdisciplinary studies. Cancer Epidemiol. Biomark. Prev., *1:* 505–513, 1992.

22. Gonzalez, F. J. Genetic polymorphism and cancer susceptibility: fourteenth Sapporo Cancer Seminar. Cancer Res., *55:* 710–715, 1995.

23. Kawajiri, K., Nakachi, K., Watanabe, J., and Hayashi, S. The *CYP1A1* gene and cancer susceptibility. Crit. Rev. Oncol. Hematol., *14:* 77–87, 1993.

24. Khoury, M. J., Adams, M. J., and Flanders, W. D. An epidemiological approach to ecogenetics. Am. J. Hum. Genet., *42:* 89–95, 1988.

25. Omenn, G. S. Risk assessment, pharmacogenetics and ecogenetics. Banbury Rep., *16:* 3–13, 1984.

26. Thomas, M., Kaltin, A., Labrecque, S., Pim, D., Banks, L., and Matlashewski, G. Two polymorphic variants of wild-type *p53* differ biochemically and biologically. Mol. Cell. Biol., *2:* 1092–1100, 1999.

27. Robins, J., and Pike, M. The validity of case-control studies with nonrandom selection of controls. Epidemiology, *1:* 273–284, 1990.

28. Wang, Y. C., Chen, C. Y., Chen, S. K., Chang, Y. Y., and Lin, P. *p53* codon 72 polymorphism in Taiwanese lung cancer patients: association with lung cancer susceptibility and prognosis. Clin. Cancer Res., *5:* 129–134, 1999.

29. London, S. J., Daly, A. K., Cooper, J., Navidi, W. C., Carpenter, C. L., and Idle, J. R. Polymorphism of glutathione *S*-transferase M1 and lung cancer risk among African-Americans and Caucasians in Los Angeles County, California. J. Natl. Cancer Inst., *85:* 1246–1252, 1995.

30. Petersen, D. D., McKinney, C. E., Ikeya, K., Smith, H. H., Bale, A. E., McBride, D. W., and Nebert, D. W. Human *CYP1A1* gene: cosegregation of the enzyme inducibility phenotype and an RFLP. Am. J. Hum. Genet., *48:* 720–725, 1991.