Repair of Tobacco Carcinogen-Induced DNA Adducts and Lung Cancer Risk: a Molecular Epidemiologic Study

Qingyi Wei, Lie Cheng, Christopher I. Amos, Li-E. Wang, Zhaozheng Guo, Waun K. Hong, Margaret R. Spitz

Background: **Only a fraction of cigarette smokers develop lung cancer, suggesting that people differ in their susceptibility to this disease. We investigated whether differences in DNA repair capacity (DRC) for repairing tobacco carcinogen-induced DNA damage are associated with differential susceptibility to lung cancer.** ***Methods:*** **From August 1, 1995, through April 30, 1999, we conducted a hospital-based, case–control study of 316 newly diagnosed lung cancer patients and 316 cancer-free control subjects matched on age, sex, and smoking status. DRC was measured in cultured lymphocytes with the use of the host-cell reactivation assay with a reporter gene damaged by a known activated tobacco carcinogen, benzo[*a*]pyrene diol epoxide. Statistical tests were two-sided.** ***Results:*** **Overall, lower DRC was observed in case patients than in control subjects ($P<.001$) and was associated with a greater than twofold increased risk of lung cancer. Compared with the highest DRC quartile in the control subjects and after adjustment for age, sex, pack-years of smoking, family history of cancer, and other covariates, reduced DRC was associated with increased risk of lung cancer in a dose-dependent fashion (odds ratio [OR] = 1.8 with 95% confidence interval [CI] = 1.1–3.1, OR = 2.0 with 95% CI = 1.2–3.4, and OR = 4.3 with 95% CI = 2.6–7.2 for the second, third, and fourth quartiles, respectively; $P_{trend}<.001$). Case patients who were younger at diagnosis (<60 years old), female, or lighter smokers or who reported a family history of cancer exhibited the lowest DRC and the highest lung cancer risk among their subgroups, suggesting that these subgroups may be especially susceptible to lung cancer.** ***Conclusion:*** **The results provide evidence that low DRC is associated with increased risk of lung cancer. The findings from this hospital-based, case–control study should be validated in prospective studies. [J Natl Cancer Inst 2000;92:1764–72]**

Differential susceptibility to carcinogenesis is suggested by the fact that only a fraction of cigarette smokers develop smoking-related lung cancer *(1).* This variation has been suggested to be due, in part, to genetically determined variation in carcinogen metabolism *(2)* and/or variability in DNA repair capacity (DRC) *(3).* Familial aggregation of lung cancer *(4)* provides indirect evidence for heritable susceptibility in the etiology of lung cancer.

Benzo[*a*]pyrene [B(a)P], a polycyclic aromatic hydrocarbon (PAH) compound, is a classic DNA-damaging carcinogen found in tobacco smoke and in the environment as a result of fuel combustion *(5).* B(a)P bioactivation *in vivo* by cytochrome P450 and epoxide hydrolase generates highly toxic electrophilic and free-radical reactive intermediates, such as B(a)P diol epoxide (BPDE). These compounds can irreversibly damage DNA by forming DNA adducts through covalent binding or oxidation *(6,7).*

Evidence from molecular epidemiologic studies *(8,9)* indicates that smokers have higher PAH–DNA adduct levels than nonsmokers. The levels of DNA adducts detected in nonneoplastic surgical lung parenchymal samples *(10)* and in alveolar macrophages *(11)* are also higher in smokers than in former smokers, suggesting that tobacco exposure is a source of B(a)P. BPDE–DNA adducts can block the transcription of an essential gene *(12)* if they are not repaired efficiently by the nucleotide excision repair (NER) pathway *(13).* BPDE–DNA adducts are repaired more efficiently in the transcribed than in the untranscribed strand *(14,15);* thus, mutations occur more frequently in the untranscribed strand. In addition, BPDE has been found to bind preferentially to mutational hot spots, resulting in G to T transversions, in the p53 tumor suppressor gene in cells that are deficient in NER *(16).* Therefore, a high frequency of G → T transversion mutations in the p53 gene in tobacco-related cancers, including lung cancer *(17,18),* strongly suggests an etiologic link between exposure to tobacco carcinogens such as B(a)P, inefficient DRC, and lung cancer risk.

Low DRC in surrogate tissue, such as lymphocytes, has been linked to elevated risk for skin cancer *(19,20).* Because the levels of smoking-induced DNA adducts in lymphocytes and in lung tissue are associated *(21),* lymphocytes should be a relevant and accessible surrogate tissue for estimating DRC in lung tissue. Using cultured lymphocytes, we previously conducted a pilot case–control study of 51 lung cancer patients and 56 healthy control subjects *(22)* and reported a fivefold increased risk for lung cancer associated with low DRC. This preliminary finding encouraged us to conduct a larger molecular epidemiologic study. In this article, we describe the analysis of DRC data obtained using the host-cell reactivation (HCR) assay for 316 patients with lung cancer and 316 matched healthy control subjects not included in our previously published pilot study *(22).*

Subjects and Methods

Study Subjects

The case patients were recruited from The University of Texas M. D. Anderson Cancer Center, Houston, during the period from August 1, 1995, through April 30, 1999. All patients with newly diagnosed, histopathologically confirmed lung cancer were eligible for the study. Healthy control subjects were recruited from Kelsey Seybold Clinics, a local managed care organization with multiple clinics throughout the Houston metropolitan area. The potential control subjects were first surveyed by a short questionnaire for willingness to partici-

Affiliations of authors: Q. Wei, L. Cheng, C. I. Amos, L.-E. Wang, Z. Guo, M. R. Spitz (Department of Epidemiology, Division of Cancer Prevention), W. K. Hong (Department of Thoracic and Head and Neck Medical Oncology), The University of Texas M. D. Anderson Cancer Center, Houston.

Correspondence to: Qingyi Wei, M.D., Ph.D., Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Box 189, 1515 Holcombe Blvd., Houston, TX 77030 (e-mail: qwei@mdanderson.org).

See "Notes" following "References."

pate in research studies and to provide preliminary data on matching characteristics, such as smoking behavior and demographics *(23).* A computer database of more than 60 000 potential control subjects was thus constructed and used to identify eligible control subjects, who were selected to match the case patients by age (±5 years), sex, ethnicity, and smoking status (never, former, or current smokers). "Smokers" were defined as those who had smoked more than 100 cigarettes in their lifetimes; they were further divided into "former smokers," who had quit smoking more than 1 year previously, and "current smokers." Pack-years were calculated from duration and amount of smoking.

Once an eligible case patient was recruited, a control subject with the closest match on these variables was then identified and contacted. If this control subject refused or was not able to participate, the next eligible control subject would be contacted. Because the parent study has not yet been completed and suitably matched control subjects were not available for all case patients, we relaxed the matching criteria for smoking status (i.e., ever or never) in this analysis. Each eligible subject was then scheduled for an interview. After written informed consent was obtained from the subjects, an interviewer administered a structured questionnaire to collect detailed information on demographic data and lung cancer risk factors, such as smoking status and family history of first-degree relatives with cancer. The exclusion criteria were prior chemotherapy or radiotherapy (for the case patients) and prior cancer and any recent blood transfusion (for all subjects). The response rate for participation was 77.4% among case patients and 73.3% among control subjects, a difference that was not statistically significant ($P = .254$). The study protocol was approved by the institutional review boards of The University of Texas M. D. Anderson Cancer Center and the Kelsey Seybold Clinics.

Blood Sample Processing and Lymphoblastoid Cell Lines

Each subject donated a 20-mL blood sample that was drawn into a heparinized VACUTAINER tube (Becton Dickinson, Franklin Lakes, NJ). Within 24 hours, lymphocytes were isolated by Ficoll gradient centrifugation at approximately 1200*g* for 35 minutes at room temperature and subsequently frozen in medium containing 50% fetal bovine serum, 40% RPMI-1640 medium, and 10% dimethyl sulfoxide. The cells were stored in a −80 °C freezer in 1.5-mL aliquots until they were thawed for the assays. Four Epstein-Barr virus-immortalized human lymphoblastoid cell lines from the Human Genetic Mutant Cell Repositories (Camden, NJ) were used as the assay controls: two apparently normal (i.e., repair-proficient) cell lines (GM00892B and GM00131A) and two repair-deficient xeroderma pigmentosum (XP) cell lines (GM02345B: XP-A and GM02246B: XP-C). These cell lines were used to monitor the assay conditions before and during the study.

Preparation of Plasmids With BPDE-Induced Adducts

We selected BPDE as the test agent from among many other tobacco carcinogens, both to be consistent with previous studies *(16,22,24)* and because BPDE–DNA adducts are stable. BPDE (National Cancer Institute #L0137; >98% pure) was purchased from Midwest Research Institute (Kansas City, MO) as a white powder and was completely dissolved in tetrahydrofuran (Sigma Chemical Co., St. Louis, MO). The plasmid used for these experiments was pCMV*cat,* which was a gift from Dr. Lawrence Grossman, The Johns Hopkins University, Baltimore, MD. pCMV*cat* contains the bacterial chloramphenicol acetyltransferase (CAT) gene as a reporter gene and an enhancer–promoter unit of human cytomegalovirus *(25).* Plasmids were treated with BPDE with the use of a modification of a previously described method *(22,24).* Briefly, purified plasmid was dissolved in Tris-EDTA buffer (pH 7.8) at a concentration of 500 μg/mL. Aliquots of plasmid solution (1 mL) were placed in microcentrifuge tubes, dissolved BPDE was added to each tube to a final concentration of 60 μ*M,* and the mixtures were incubated for 3 hours in a dark room. These treatments were performed in one batch. After BPDE treatment, plasmids were precipitated three times with 70% ethanol to remove free BPDE, dissolved in Tris-EDTA buffer at a final concentration of 50 μg/mL, and stored in aliquots in a −20 °C freezer. Previous experiments *(22,24)* have shown that, at a final concentration of 60 μ*M,* BPDE induces at least one adduct per plasmid and does not cause conformational changes in the plasmids, as confirmed by 0.8% agarose gel electrophoresis after treatment. These repurified, BPDE-treated plasmids were used for all HCR assays reported in this article.

Lymphocyte Transfections With Plasmids

The frozen lymphocyte samples were coded so that laboratory personnel were blinded to status (i.e., whether they were from case patients or from control subjects). When enough cell samples had accumulated to perform the repair assays in batches (about 10–20 samples from approximately equal numbers of case patients and control subjects), the frozen cell samples were thawed and processed as described previously *(22,24).* Briefly, the cells in each cryogenic vial (1.5 mL) were quickly thawed and mixed (before the last trace of ice had disappeared) with 8.5 mL of thawing medium (50% fetal bovine serum, 40% RPMI-1640 medium, and 10% dextrose), which ensured a cellular viability of more than 80% as tested by 0.4% trypan blue dye (Sigma Chemical Co.) exclusion test *(26),* and then washed with the thawing medium. The cells were next stimulated so that they would take up the plasmids *(24)* and exhibit the NER activity *(27,28)* that removes the BPDE-induced DNA adducts. For stimulation, cells were incubated at 37 °C for 72 hours in RPMI-1640 medium supplemented with 20% fetal bovine serum and 56.25 μg/mL phytohemagglutinin (Murex Diagnostics, Norcross, GA). After stimulation, the number of viable, large lymphoblasts in the culture for each sample was counted to calculate the blastogenic rate. The lymphoblasts from each subject were then divided into four aliquots, each containing approximately 2×10^6 cells, for duplicate transfections with untreated plasmids (the baseline for comparison) and duplicate transfections with BPDE-treated plasmids. The transfections were performed by the diethylaminoethyl-dextran (Pharmacia Biotech Inc., Piscataway, NJ) method *(29).*

HCR Assay

The HCR assay measures the activity of the CAT gene, a bacterial drug resistance gene, in cells that have been transfected with BPDE-treated plasmid *(24).* Because a single unrepaired DNA adduct can effectively block CAT transcription *(30),* any CAT activity will reflect the ability of the transfected cells to remove BPDE-induced adducts from the plasmids. Therefore, this assay provides a quantitative measurement of the DRC of the host cells.

CAT activity in transfected cells was measured as described previously *(24).* Briefly, the tubes containing the cell culture with transfections were centrifuged at approximately 800 rpm for 10 minutes at room temperature; the cell pellets were collected and washed twice with 1.5 mL of Tris-buffered saline (TBS) and resuspended in 31.5 μL of 0.25 *M* Tris-TBS in a 1.5-mL tube. The cells were lysed by three 10-minute cycles of freezing and thawing in a dry ice–ethanol bath and a 37 °C water bath. Cell extracts were then assayed for CAT expression or activity. The activity of the repaired CAT gene was measured by a scintillation counter for the formation of [^{3}H]monoacetylated and [^{3}H]diacetylated chloramphenicols through the reaction between chloramphenicol and [^{3}H]acetyl coenzyme A catalyzed by CAT protein in the cell extract. DRC is defined as the ratio of the CAT activity of cells transfected with BPDE-treated plasmids to that of cells transfected with untreated plasmids—i.e., $DRC = (CAT_{BPDE60}/CAT_{BPDE0}) \times 100\%$. The CAT activity of cells transfected with undamaged plasmids provides an experimental internal control because it is derived under the same experimental conditions as the CAT_{BPDE60} *(24)* and from the same number of cells from the same individual. With the use of this assay, the DRC of the normal control cell lines was approximately 20% (GM00829B) and 15% (BM00131A) because the measured level of DRC is dependent on the dose of BPDE (60 μ*M*) used to damage the plasmids *(24).* At the same BPDE dose, the DRC of the two XP cell lines was less than 1% (GM02345B: XP-A) and approximately 1% (GM02246B: XP-C); therefore, this BPDE dose can differentiate cells that are repair proficient or repair deficient.

Statistical Analysis

The distributions of matching variables, including age, sex, ethnicity (only Caucasians), smoking status, and pack-years of smoking were first examined for the adequacy of the matching procedure. DRC was analyzed as a continuous variable before and after natural logarithmic transformation. Student's *t* test was used to compare the differences in DRC between groups. Whenever the variance of the groups varied significantly, Student's *t* tests with unequal variances were used for comparisons. Correlation analyses were performed for DRC and selected variables. The blastogenic rate (percentage of lymphocytes that responded to phytohemagglutinin stimulation), the actual amount of radioactivity of ^{3}H-labeled acetyl base that was transferred by CAT in cells transfected with untreated plasmids (i.e., the baseline CAT expression level, in counts per minute [cpm]), and cell storage time (in months) were also recorded for comparison between case patients and control subjects.

For calculation of crude odds ratios (ORs) and confidence intervals (CIs), the median DRC of control subjects was used as the cutoff value: Values greater than this median were considered to be high (i.e., efficient) DRC, and values

below the median were considered to be low (i.e., suboptimal) DRC. The number of first-degree relatives with cancer was included as both continuous and dichotomized (i.e., yes and no) variables. For logistic regression analysis, dummy variables of the median and quartile DRC were created to calculate the ORs and 95% CIs. Finally, adjusted ORs were calculated by fitting unconditional multivariate logistic regression models with adjustment for age, sex, pack-years of smoking, family history of cancer, blastogenic rate, cell storage time, and baseline CAT activity. To perform the linear trend test, the quartile dummy variables were recoded as one continuous variable (1–4 for the lowest to highest quartiles) and fitted into a logistic regression model with and without adjustment for covariates. All statistical tests were two-sided and were performed with Statistical Analysis System software (version 6; SAS Institute Inc., Cary, NC).

RESULTS

From August 1, 1995, through April 30, 1999, a total of 381 eligible patients with newly diagnosed and histopathologically confirmed lung cancer without prior treatment were consecutively recruited at The University of Texas M. D. Anderson Cancer Center. These patients had donated blood samples that provided sufficient numbers of viable lymphocytes for the HCR assay. Because relatively small numbers of African-American and Mexican-American patients have been recruited thus far, we analyzed only the 316 non-Hispanic Caucasians. Their tumor types included adenocarcinomas (46%), small-cell carcinomas (10%), squamous cell carcinomas (20%), non-small-cell carcinomas (13%), and carcinomas of other histopathologic types (including large-cell and giant-cell carcinomas) (11%). During the same time period, 316 healthy non-Hispanic Caucasian subjects who had no history of malignancies were selected from the available control subject database by the matching criteria.

The distribution of the matching variables between case patients and control subjects was first examined to determine whether matching was adequate. As shown in Table 1, case patients and control subjects were well matched in their distribution into age groups of less than 60 years old, 60–69 years old, and 70 years old or older ($P = .832$) and on sex ($P = 1.000$). The mean age was 61.3 years (±9.5 years; range, 34–84 years) for case patients and 61.2 years (±10.6 years; range, 32–86 years) for control subjects (Table 2). Although there were more current smokers among case patients (46%) than among control subjects (37%) and more former smokers among control subjects (52%) than among case patients (45%) (Table 1), the difference in smoking status did not reach statistical significance ($P = .065$). Furthermore, the mean pack-years of smoking was only slightly higher in case patients (53.2 years) than in control

Table 1. Distribution of select characteristics and DNA repair capacity (DRC) between case patients and healthy control subjects

Variable	Case patients (n = 316)		Control subjects (n = 316)			
	No. (%)	DRC,* mean ± SD†	No. (%)	DRC, mean ± SD	P‡	P§
Age, y					.832	
<60	128 (41)	7.6 ± 2.8	123 (39)	9.8 ± 5.0		<.001
60–69	123 (39)	8.2 ± 3.2	122 (39)	9.9 ± 5.0		.002
⩾70	65 (21)	8.5 ± 3.2	71 (22)	9.7 ± 3.4		<.001
Trend test‖		$P = .050$		$P = .994$		
Sex					1.000	
Male	170 (54)	8.6 ± 3.0	170 (54)	10.3 ± 5.1		<.001
Female	146 (46)	7.4 ± 3.0	146 (46)	9.3 ± 4.1		<.001
Student's *t* test¶		$P<.001$		$P = .058$		
Smoking status					.065	
Never	28 (9)	7.1 ± 3.3	33 (10)	9.0 ± 4.2		.059
Former≠	141 (45)	8.3 ± 3.0	165 (52)	9.6 ± 4.1		.002
Current	147 (47)	8.0 ± 3.1	118 (37)	10.3 ± 5.4		<.001
Trend test‖		$P = .508$		$P = .081$		
Smoked in last 24 h	53 (17)	8.1 ± 3.6	102 (32)	10.1 ± 5.4		.007
Student's *t* test**		$P = .225$		$P = .287$		
Pack-years of smoking††					.052	
0–19	46 (15)	7.5 ± 3.3	61 (19)	9.8 ± 4.5		.005
20–39	59 (19)	7.9 ± 2.4	78 (25)	9.4 ± 4.7		.016
40–60	98 (31)	7.8 ± 2.9	83 (26)	10.0 ± 5.4		.001
>60	113 (36)	8.6 ± 3.4	94 (30)	10.1 ± 4.1		.005
Trend test‖		$P = .066$		$P = .550$		
No. of first-degree relatives with any cancer					.025	
None	98 (31)	8.3 ± 3.1	127 (40)	9.1 ± 3.8		.094
1	125 (40)	8.0 ± 2.9	120 (38)	10.4 ± 5.4		<.001
>1	93 (29)	8.0 ± 3.4	69 (22)	10.2 ± 4.9		<.001
Trend test‖		$P = .489$		$P = .044$		

*DRC was measured as the ratio of reporter gene activity in cells transfected with damaged plasmids to that in cells transfected with undamaged plasmids (×100%). *See* the "Subjects and Methods" section for details.

†SD = standard deviation.

‡Two-sided chi-square tests for differences in the distributions of the numbers of each category of the subgroup between case patients and control subjects.

§Two-sided Student's *t* tests for the differences in DRC between case patients and control subjects.

‖Calculated in the general linear models with the use of natural log-transformed DRC values.

¶*P* value for comparison between males and females.

≠Including some recent quitters. Former smokers were those who quit smoking between 4 and 12 months and recent quitters (50 case patients and eight control subjects) were those who quit smoking in the last 4 weeks before diagnosis.

***P* value for comparison of people who smoked within the last 24 hours only with never smokers.

††Defined as number of packs of cigarettes smoked per day × number of years of smoking.

Table 2. Comparison of select variables and DNA repair capacity (DRC) covariates between case patients and control subjects*

Variable	Mean ± SD		P†
	Case patients (n = 316)	Control subjects (n = 316)	
Age, y	61.3 ± 9.5	61.2 ± 10.6	.887
Pack-years of smoking	53.2 ± 33.7	49.2 ± 36.3	.148
Blastogenic rate, %‡	32.9 ± 23.3	34.2 ± 23.5	.469
Cell storage time, mo	8.4 ± 6.5	5.9 ± 4.4	<.001
Baseline CAT activity, cpm§	31 349 ± 25 449	28 492 ± 20 406	.120
DRC, %‖	8.1 ± 3.1	9.8 ± 4.7	<.001
Natural log-transformed DRC, %	2.0 ± 0.4	2.2 ± 0.4	<.001

*SD = standard deviation; CAT = chloramphenicol acetyltransferase; cpm = counts per minute.

†Two-sided Student's *t* tests for differences between case patients and control subjects.

‡Percentage of cells that were stimulated by phytohemagglutinin.

§CAT expression from a mock transfection was approximately 100 cpm.

‖Measured as the ratio of CAT activity in cells transfected with damaged plasmids to that in cells transfected with undamaged plasmids (×100%). *See* the "Subjects and Methods" section for details.

subjects (49.2 years), and this difference was not statistically significant ($P = .148$) (Table 2), suggesting that the matching on smoking status was also adequate. More case patients (69%) than control subjects (60%) reported having a first-degree relative with any cancer ($P = .025$) (Table 1). However, these variables (age, sex, pack-years of smoking, and family history of cancer) were further adjusted for in the multivariate logistic regression analyses.

When DRC was analyzed as a continuous variable (Table 2), the mean DRC was 8.1% in case patients and 9.8% in control subjects, representing an average 17% lower DRC in case patients ($P<.001$). This difference remained statistically significant after natural log transformation of the data ($P<.001$; Table 2). Although non-log-transformed DRCs were somewhat skewed to high values, the results without transformation were qualitatively similar to those with natural log transformation. Therefore, we used untransformed DRCs in the rest of the analyses. There was a more than fivefold variation in DRC (range, 3%–15%) among the 316 control subjects (intervariation), representing a coefficient of variation of nearly 48% as compared with the previously reported coefficient of variation of 15% for repeated assays of healthy volunteers (intravariation) over a 6-week period *(24)*. This difference suggests that the main source of the variation in DRC was individual variation and that the comparison using one cross-sectional measurement of DRC was adequate for this study.

Further stratification of DRC revealed that all subgroups of case patients consistently exhibited significantly lower mean DRC than did the control subjects (Table 1). Patients who were younger at lung cancer diagnosis (<60 years old) had the lowest mean DRC. Compared with men, women had lower mean DRC, among both case patients ($P<.001$) and control subjects ($P = .058$). Case patients who were nonsmokers exhibited the lowest mean DRC, although the difference did not reach statistical significance ($P = .059$) because of the small numbers in this stratum. Among control subjects, current and heaviest smokers (>60 pack-years) exhibited the highest DRC compared with former, lighter (<60 pack-years), and never smokers, but none of the differences were statistically significant. However, there were no differences in DRC between former and current smokers in either case patients or control subjects.

Because there was a nonstatistically significant trend of increasing DRC in both case patients and control subjects with increasing smoking intensity as measured by pack-years (Table 1), we assessed the effect of smoking on DRC. Among current smokers, 36% (53 of 147) of case patients and 86% (102 of 118) of control subjects reported smoking within the 24 hours before the blood samples were drawn, suggesting that case patients may have reduced smoking because of illness. However, the mean DRC of these recent smokers was not different from that of all current, former, or never smokers (Table 1), suggesting that recent smoking does not have an immediate impact on DRC.

Self-reported family history of cancer, regardless of the number of first-degree relatives with any cancer (except for nonmelanoma skin cancer), was associated with lower DRC in case patients and with higher DRC in control subjects (Table 1), but there was no statistical evidence of interaction (data not shown).

So that we could further explore the statistically significant sex difference in DRC, DRC was further stratified by sex and other select variables (Table 3). There were no age differences between case patients and control subjects for both sexes, but there were significantly more heavy smokers (i.e., those who had smoked ⩾40 pack-years) among male case patients (74%) than among female case patients (54%) ($P<.001$) or male control subjects (62%) ($P = .027$). Moreover, statistically significantly more female case patients (72%) than female control subjects (58%) reported a first-degree relative with any cancer ($P = .015$). However, statistically significant differences in DRC between case patients and control subjects remained in these subgroups stratified by sex (Table 3), suggesting that the lower DRC in female case patients as compared with male case patients was not due to confounding by age, smoking, or family history of cancer.

To determine whether the treatment of the cells could have had different effects on the cells from case patients and those from control subjects, we evaluated the effects of freezing and storage by comparing blastogenic rates, cell storage time, and baseline CAT expression in transfected cells from case patients and control subjects. The blastogenic rates of cells from case patients (32.9%) and control subjects (34.2%) were almost identical (Table 2), suggesting that the effects of freezing and storage on cellular response to phytohemagglutinin and cell growth were similar for case patients and control subjects. The average baseline CAT activity in cells from all subjects that were transfected with untreated plasmid (approximately 30 000 cpm) was nearly 300 times greater than that of background CAT activity (about

Table 3. Comparison of DNA repair capacity (DRC) by sex and select variables between case patients and control subjects

Variable	Case patients		Control subjects			
	No. (%)	DRC, mean ± SD*	No. (%)	DRC, mean ± SD*	P†	P‡
Male	170 (100)		170 (100)			
Age, y					1.000	
<60	63 (37)	8.5 ± 2.9	63 (37)	10.7 ± 6.3		.013
⩾60	107 (63)	8.7 ± 3.2	107 (63)	10.0 ± 4.3		.013
Student's t test§		P = .642		P = .446		
Pack-years of smoking‖					.027	
<40	45 (26)	8.4 ± 3.0	64 (38)	10.6 ± 5.6		.018
⩾40	125 (74)	8.7 ± 3.1	106 (62)	10.0 ± 4.8		.015
Student's t test†		P = .598		P = .450		
No. of first-degree relatives with cancer					.308	
No	56 (33)	9.0 ± 3.1	65 (38)	9.0 ± 3.5		.997
Yes	114 (67)	8.4 ± 3.0	105 (62)	11.0 ± 5.8		<.001
Student's t test§		P = .203		P = .015		
Female	146 (100)		146 (100)			
Age, y					.554	
<60	65 (45)	6.8 ± 2.5	60 (41)	8.8 ± 3.0		<.001
⩾60	81 (55)	7.8 ± 3.3	86 (59)	9.6 ± 4.6		.003
Student's t test§		P = .045		P = .169		
Pack-years of smoking‖					.160	
<40	67 (46)	7.2 ± 3.6	79 (54)	8.7 ± 3.1		.002
⩾40	79 (54)	7.5 ± 2.3	67 (46)	9.9 ± 4.8		<.001
Student's t test§		P = .623		P = .071		
No. of first-degree relatives with cancer					.015	
No	42 (29)	7.2 ± 2.5	62 (42)	9.0 ± 4.2		.006
Yes	104 (71)	7.4 ± 3.2	84 (58)	9.5 ± 3.9		<.001
Student't t test§		P = .630		P = .515		

*DRC was measured as the ratio of reporter gene activity in cells transfected with damaged plasmids to that in cells transfected with undamaged plasmids (×100%). *See* the "Subjects and Methods" section for details. SD = standard deviation.

†Two-sided chi-square tests for differences in the distributions of the numbers of each category of the subgroup between case patients and control subjects.

‡Two-sided Student's t test for the differences in DRC between case patients and control subjects.

§P value for comparison within case patients or control subjects.

‖Defined as number of packs of cigarettes smoked per day × number of years of smoking.

100 cpm for untransfected cells) (data not shown), suggesting that the transfections were successful. Cell storage time was longer for case patients (average, 8.4 months) than for control subjects (average, 5.9 months) because control subject selection, of necessity, lagged behind case patient recruitment. However, when the analysis was restricted to the 161 case patients and 247 control subjects whose samples were stored for fewer than 10 months so that there was no difference in average storage time between case patients and control subjects, the significant differences in DRC between case patients and control subjects remained (data not shown). This result suggests that the storage time had no effect on DRC. Case patients had nonstatistically significantly higher baseline CAT expression levels than did control subjects, implying that the disease status and longer cell storage time did not have adverse effects on transfection rate and cell growth in the case patients. Furthermore, none of the assay parameters that could have differed between case patients and control subjects (blastogenic rate, cell storage time, and baseline CAT expression) was correlated with DRC (r = –.058, P = .144; r = .045, P = .261; and r = .025, P = .516, respectively). Nevertheless, these variables were also adjusted for in subsequent multivariate logistic regression analysis.

To assess whether DRC was associated with the histologic type of the tumor, we further grouped the case patients by histologic type and then compared the DRC of each type with that of the control subjects. Patients with each histologic type had statistically significantly lower DRC than did the control subjects (data not shown), except for patients with small-cell carcinomas (P = .056), probably because of the small number of patients in this category (n = 32) (data not shown). There were no differences in the DRC among case patients by histologic type (data not shown), a finding that is consistent with DRC being a constitutional rather than a tumor marker.

The effect of DRC on risk for lung cancer was evaluated by logistic regression analysis. As shown in Table 4, four models with different methods of categorizing DRC were tested with the use of the maximum likelihood method. When the DRC and its natural log-transformed values were used in the logistic regression model with adjustment for age, sex, pack-years of smoking, blastogenic rate, cell storage time, baseline CAT expression, and family history of cancer, the crude and adjusted ORs (per DRC unit decrease) for both untransformed and natural log-transformed DRC values were statistically significantly elevated (OR = 1.1 [95% CI = 1.1–1.2] and OR = 3.1 [95% CI = 2.0–4.6], respectively; Table 4). When DRC values were dichotomized by the median DRC of the control subjects (with DRC above the median characterized as high and that below the median as low), the crude OR associated with low DRC was 1.9 (95% CI = 1.4–2.6). After adjustment for age, sex, pack-years of smoking, blastogenic rate, cell storage time, baseline CAT expression, and family history of cancer, the OR associated with low DRC was 2.1 (95% CI = 1.5–3.0) (Table 4).

When the DRC values were further divided by quartile of DRC in control subjects, it was again evident that decreased DRC was associated with increased risk. By use of the highest quartile of the DRC as the reference group, the crude ORs for

Table 4. Logistic regression analysis of DNA repair capacity (DRC) in case patients and control subjects*

Variable	Case patients		Control subjects		Crude OR (95% CI)	Adjusted† OR (95% CI)
	No.	%	No.	%		
DRC (%)‡	316	100	316	100	1.1 (1.1–1.2)	1.2 (1.1–1.2)
Log-transformed DRC (%)§	316	100	316	100	3.1 (2.0–4.6)	4.0 (2.6–6.3)
DRC (%; by median)‖						
⩾8.8	110	35	158	50	1.0 (referent)	1.0 (referent)
<8.8	206	65	158	50	1.9 (1.4–2.6)	2.1 (1.5–3.0)
DRC (%; by quartiles)¶						
>11.5	38	12	76	24	1.0 (referent)	1.0 (referent)
8.8–11.5	72	23	82	26	1.8 (1.1–2.9)	1.8 (1.1–3.1)
7.0–8.7	77	24	81	26	1.9 (1.2–3.1)	2.0 (1.2–3.4)
<7.0	129	41	77	24	3.4 (2.1–5.4)	4.3 (2.6–7.2)
Trend test#					*P*<.001	*P*<.001

*OR = odds ratio; CI = confidence interval.

†Adjusted in a logistic regression including age (years), sex, baseline chloramphenicol acetyltransferase (CAT) activity, blastogenic rate (%), cell storage time (months), pack-years of smoking, and family history of first-degree relatives with any cancer.

‡DRC was measured as the ratio of CAT activity in cells transfected with damaged plasmids to that in cells transfected with undamaged plasmids (×100%). *See* the "Subjects and Methods" section for details. DRC (%) was fitted in the logistic regression model as a continuous variable.

§Natural log-transformed DRC values were fitted in the logistic regression model as a continuous variable.

‖Median of the control subjects' values.

¶Quartiles of the control subjects' values.

#Adjusted for the same variables as described above.

DRC values lower than values in the 75th, 50th, and 25th quartiles were 1.8 (95% CI = 1.1–2.9), 1.9 (95% CI = 1.2–3.1), and 3.4 (95% CI = 2.1–5.4), respectively. The ORs after adjustment for age, sex, pack-years of smoking, blastogenic rate, cell storage time, baseline CAT expression, and family history of cancer were nearly identical (OR = 1.8 [95% CI = 1.1–3.1], OR = 2.0 [95% CI = 1.2–3.4], and OR = 4.3 [95% CI = 2.6–7.2], respectively) to the crude ORs. This trend of increasing risk with decreasing DRC was statistically significant (*P*<.001) for both the crude and adjusted ORs (Table 4).

Finally, the risk of lung cancer associated with low DRC in subgroups of age, sex, pack-years of smoking, and family history was further evaluated by comparing the adjusted ORs stratified by DRC quartile as defined by the DRC values of control subjects. As shown in Fig. 1, with the use of the highest quartile (DRC >11.5%) as the referent group, those individuals of the entire 632-subject population who were in the lowest quartile (DRC <7.0%) had a more than fourfold increased risk of lung cancer; individuals under 60 years of age had an OR of 4.9 (95% CI = 2.1–11.6), females had an OR of 7.6 (95% CI = 3.2–18.0), individuals who had smoked fewer than 40 pack-years had an OR of 4.8 (95% CI = 2.1–11.0), and individuals who had a positive family history of cancer had an OR of 6.3 (95% CI = 3.2–12.1). These findings are consistent with the notion that a fraction of lung cancer patients are especially susceptible to tobacco carcinogen-induced DNA damage due to suboptimal DNA repair.

Discussion

In this large molecular epidemiologic study of DRC in lung cancer, which we believe to be the first of its kind, we demonstrated that reduced DRC had a statistically significant effect on risk of lung cancer after controlling for potential confounding variables (age, sex, pack-years of smoking, and number of first-degree relatives with cancer) and covariates (blastogenic rate of lymphocytes, cell storage time, and baseline CAT expression). Although the DRC assay might be influenced by other, unmeasured confounding factors, the relatively large sample size, matching design, and consistent differences observed in the DRC of case patients and control subjects in all of the subgroups analyzed suggest that the results are unlikely to be spurious. Rather, they strongly support the hypothesis that reduced DRC is associated with increased risk of lung cancer. The overall 2.1-fold increased risk (95% CI = 1.5–3.0) is consistent with the finding of our previous small pilot study *(22),* which reported a fivefold increased risk but with a wider 95% CI of 2.1–15.7. Furthermore, the finding of a dose-dependent association between decreased DRC and increased risk further strengthens the biologic plausibility of our hypothesis. The nearly fourfold increased risk of lung cancer among individuals, regardless of smoking status, whose DRC falls in the lowest quartile is consistent with the notion that a fraction of the population is predisposed to lung cancer *(3,19,31,32).* The most susceptible subgroups on the basis of their low DRC were case patients who were young (<60 years old), female, or light smokers or who reported a family history of cancer.

Several lines of evidence support the hypothesis that innate DRC is central in controlling multistage carcinogenesis. Because the spontaneous mutation rate is so low, a mutator phenotype may be required for multistage carcinogenesis *(33).* BPDE-induced adducts are highly mutagenic and form at a higher rate in the hotspots of the p53 tumor suppressor gene in cells that fail to remove the damage efficiently *(15,16),* resulting in the G→T hotspot mutations in p53 that are commonly seen in smoking-induced cancers *(17,18).* Although a direct link between BPDE exposure, high mutation rate, and low DRC is experimentally proven *(16),* population data that support such a link are lacking. However, our findings of low DRC in lung cancer patients are consistent with results from recent studies of DNA adducts. A lower DRC would result in higher levels of tobacco-related DNA adducts. For example, lung cancer patients were found to have significantly higher levels of aromatic DNA adducts and 7,8-dihydro-8-oxo-2′-deoxyguanosine, respectively, than control subjects *(34).* We have also demonstrated previously that

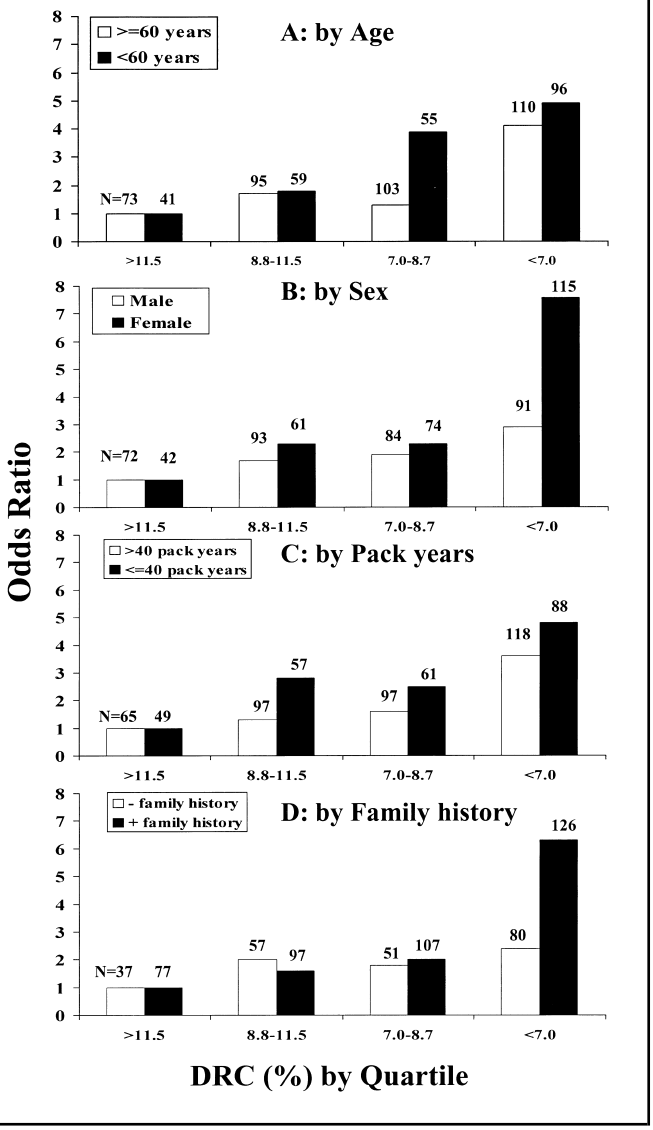

Fig. 1. Graphical presentation of odds ratios (ORs) for developing lung cancer among all 632 participants in the study, according to the DNA repair capacity (DRC) quartiles defined by the control subjects, stratified by age, sex, smoking, and family history of cancer. ORs (with the use of the highest DRC quartile as the reference group) are presented for subgroups by **A**) age (≥60 years old versus <60 years old), **B**) sex (male versus female), **C**) pack-years of smoking (≤40 pack-years versus >40 pack-years), and **D**) family history of cancer (positive versus negative). The trend for increased risk associated with decreased DRC was statistically significant for subgroups based on age (P<.001), sex (P<.001 for females, and P = .003 for males), pack-years of smoking (P<.001), and positive family history (P<.001) but not for the subgroup based on negative family history (P = .092). The most susceptible subgroups (i.e., those in the lowest DRC quartile who were young at lung cancer diagnosis, were female, had smoked 40 pack-years or less, or had a positive family history of cancer) had the greatest increase in lung cancer risk, ranging between fivefold and eightfold.

lung cancer patients with a low DRC tended to have a higher level of BPDE–DNA adducts than individuals (including case and control subjects) who had a proficient DRC *(3).*

Our observation that women have significantly lower DRC than men is consistent with our findings in a previous study, in which women exhibited lower DRC than men and tended to have a much higher risk of nonmelanoma skin cancer if they had exposure to sunlight *(20).* This finding is also consistent with recent reports *(35,36)* that women have a higher lung cancer risk than men given similar exposure to tar. It has been suggested that postmenopausal women tend to have decreased DRC, which can be increased by supplemental estrogen *(32),* but other investigators *(37)* have reported that estrogen suppresses DRC.

There are other lines of evidence for sex differences in cancer risk. For instance, among lung cancer patients, female patients had a higher level of DNA adducts than male patients after adjustment for smoking dose *(38).* In another study *(39),* female smokers with the GSTM1 null genotype, which results in diminished glutathione *S*-transferase (GST) activity, had the greatest lung cancer risk compared with other groups of females and males with different GSTM1 genotypes. This finding also suggests that women are at a greater risk of lung cancer than are men, given equal smoking. The absence of detoxifying GST activity may result in an excess of internal exposure to tobacco carcinogens, leading to a higher level of DNA damage or adduct formation *(2).* Finally, in our study, the female case patients in the lowest DRC quartile exhibited the greatest risk of lung cancer (more than sevenfold compared with female control subjects), suggesting that factors in addition to DRC may also play a role in this subgroup. It has been reported that gastrin-releasing peptide receptor, an X-linked growth factor receptor, contributes to the excess risk of lung cancer in females *(40).* Further studies on the interaction between DRC and these X-linked factors and/or receptors may help clarify the sex difference in the risk of developing lung cancer.

We previously reported that increasing age (within the range of 20–60 years) was associated with decreasing DRC in apparently healthy subjects *(20).* Although there is still controversy over whether an age-related decline in DRC exists in the general population, accumulated evidence supports this hypothesis *(41).* Such an age-related decline in the DRC was, however, not evident in the control population of the study reported here. It should be noted that only a third of the control subjects were under age 60 years; thus, it might not have been possible to identify age-related changes in the relatively older group of subjects in our study. Nevertheless, our data showed consistently that case patients who were lighter smokers or were younger than 60 years at lung cancer diagnosis exhibited lower DRC than their matched control subjects, suggesting that other undetected confounders may also play a role in this subgroup.

Our results also showed that heavy smokers among both case patients and control subjects tended to have more proficient DRC than lighter smokers, suggesting that cigarette smoking may, in fact, stimulate DRC in response to the DNA damage caused by tobacco carcinogens. Such an adaptation would be consistent with a baseline DRC that can be mobilized upon increased demand for repair *(42,43).* An inducible enhancement of DRC may explain why no age-related decline in DRC was evident in the control subjects—i.e., DRC stayed high because of continued stimulation by smoking. The finding of an increase in DRC with smoking is also consistent with our previous finding *(44)* that, in head and neck cancer patients, DNA repair gene expression was increased in the heavy smokers among both case patients and control subjects relative to the lighter smokers. Inefficient repair response or inability to increase DRC on tobacco exposure may, therefore, lead to accumulation of genetic damage. Indeed, case patients who were lighter smokers appeared to exhibit the lowest DRC and had the highest risk of lung cancer, suggesting that a poor inducible repair response may have con-

tributed to the excess risk. This adaptation of DRC to smoking, if it exists, appears to be long term rather than transient because the effect was still present in former smokers but was not stronger in those who had smoked in the 24 hours before the blood was drawn (Table 1).

Compared with all other subgroups analyzed, case patients with a family history of cancer exhibited the lowest DRC, which is consistent with the previously reported low DRC in cancer patients with a positive family history of cancer *(20,45–47).* On the other hand, control subjects with a family history of cancer tended to have a higher DRC than control subjects without a family history of cancer. This finding seems paradoxical. However, we reported a similar finding in our skin cancer study *(20).*

Although several assays are available for measuring DRC, the HCR is the assay of choice for population studies *(24).* One of the advantages of the HCR assay for measuring DRC is that the plasmids, rather than the cells, are treated with the activated carcinogen BPDE. Consequently, the assay measures intrinsic cellular DRC that can otherwise be compromised by dose-dependent cytotoxicity *(48).* Transfection efficiency may influence the results of the HCR assay. In our study, however, both cell growth after stimulation and baseline reporter (CAT) gene expression of undamaged plasmids after transfection were similar in case patients and control subjects, implying that cells from case patients and control subjects had similar transfection efficiency. Furthermore, even repair-deficient cells, such as XP cells, do not have reduced transfection efficiency *(49).* Therefore, it is unlikely that the results were influenced by differences in cell growth, differential response to mitogen stimulation, or variation in transfection efficiency.

A potential limitation of this study is our use of lymphocytes as a surrogate tissue, because their DRC may not be proportional to that of lung epithelial cells. However, we assumed that, if DRC is genetically determined, then the DRC levels of different individuals can be compared across the same type of tissue. Another limitation of this study is that, with a case–control study, a design that is retrospective in nature, it is not possible to rule out a confounding effect from disease status of lung cancer that may result in decreased DNA repair in patients with the disease. However, our study should provide a scientific basis for future cohort studies to verify the role of DRC in disease etiology.

Although DRC may be genetically determined, as reflected in a family history of cancer and as seen in XP patients, possible epigenetic influences on DRC, such as smoking, need to be further elucidated in future studies. For example, plasma β-carotene may modulate the level of DNA adducts *(50),* and PAH–DNA adducts are inversely associated with plasma levels of retinol, β-carotene, and α-tocopherol *(51).* These results suggest that micronutrients may have a protective effect against DNA damage. Therefore, future studies should focus on the joint effects of micronutrients and DRC in relation to lung cancer risk.

In conclusion, the findings of this molecular epidemiologic study support our earlier finding in a small pilot study *(22)* and provide compelling evidence that reduced DRC, whether genetically or epigenetically determined, plays a role in smoking-related lung cancer. The results suggest that subgroups of patients who were younger at lung cancer diagnosis, female, or lighter smokers or who had a family history of cancer were especially likely to exhibit suboptimal DNA repair. These susceptible individuals may be identifiable on the basis of their reduced DRC. Although the DRC assay may be a useful tool for identifying such susceptible individuals, the assay precision needs to be improved, and a single marker is unlikely to be sufficiently predictive of risk. To further refine our risk assessment, we plan to apply a panel of biomarkers for genetic susceptibility, including other DNA repair phenotypes *(22,52,53),* risk genotypes of metabolic enzymes *(3),* and DNA repair enzymes *(54).*

REFERENCES

(1) Mattson ME, Pollack ES, Cullen JW. What are the odds that smoking will kill you? Am J Public Health 1987;77:425–31.

(2) Caporaso N, Landi MT, Vineis P. Relevance of metabolic polymorphisms to human carcinogenesis: evaluation of epidemiologic evidence. Pharmacogenetics 1991;1:4–19.

(3) Wei Q, Spitz MR. The role of DNA repair capacity in susceptibility to lung cancer: a review. Cancer Metastasis Rev 1997;16:295–307.

(4) Sellers TA, Bailey-Wilson JE, Elston RC, Wilson AF, Elston GZ, Ooi WL, et al. Evidence for mendelian inheritance in the pathogenesis of lung cancer. J Natl Cancer Inst 1990;82:1272–9.

(5) Phillips DH. Fifty years of benzo(*a*)pyrene. Nature 1983;303:468–72.

(6) Gelboin HV. Benzo[α]pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes. Physiol Rev 1980;60:1107–66.

(7) MacLeod MC, Tang MS. Interaction of benzo(*a*)pyrene diol-epoxides with linear and supercoiled DNA. Cancer Res 1985;45:51–6.

(8) Everson RB, Randerath E, Santella RM, Cefalo RC, Avitts TA, Randerath K. Detection of smoking-related covalent DNA adducts in human placenta. Science 1986;231:54–7.

(9) Perera FP, Santella RM, Brenner D, Poirier MC, Munshi AA, Fischman HK, et al. DNA adducts, protein adducts, and sister chromatid exchange in cigarette smokers and nonsmokers. J Natl Cancer Inst 1987;79:449–56.

(10) Geneste O, Camus AM, Castegnaro M, Petruzzelli S, Macchiarini P, Angeletti CA, et al. Comparison of pulmonary DNA adduct levels, measured by 32P-postlabelling and aryl hydrocarbon hydroxylase activity in lung parenchyma of smokers and ex-smokers. Carcinogenesis 1991;12:1301–5.

(11) Izzotti A, Rossi GA, Bagnasco M, De Flora S. Benzo[*a*]pyrene diolepoxide–DNA adducts in alveolar macrophages of smokers. Carcinogenesis 1991;12:1281–5.

(12) Tang MS, Pierce JR, Doisy RP, Nazimiec ME, MacLeod MC. Differences and similarities in the repair of two benzo[*a*]pyrene diol epoxide isomers induced DNA adducts by uvrA, uvrB, and uvrC gene products. Biochemistry 1992;31:8429–36.

(13) Sancer A. DNA repair in humans. Annu Rev Genet 1995;29:69–105.

(14) Chen RH, Maher VM, Brouwer J, van de Putte P, McCormick JJ. Preferential repair and strand-specific repair of benzo[*a*]pyrene diol epoxide adducts in the HPRT gene of diploid human fibroblasts. Proc Natl Acad Sci U S A 1992;89:5413–7.

(15) Denissenko MF, Pao A, Pfeifer GP, Tang M. Slow repair of bulky DNA adducts along the nontranscribed strand of the human p53 gene may explain the strand bias of transversion mutations in cancers. Oncogene 1998; 16:1241–7.

(16) Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[*a*]pyrene adducts at lung cancer mutational hotspots in P53. Science 1996;274:430–2.

(17) Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. Science 1991;253:49–53.

(18) Bennett WP, Hussain SP, Vahakangas KH, Khan MA, Shields PG, Harris CC. Molecular epidemiology of human cancer risk: gene–environment interactions and p53 mutation spectrum in human lung cancer. J Pathol 1999; 187:8–18.

(19) Kraemer KH, Lee MM, Andrews AD, Lambert WC. The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. Arch Dermatol 1994;130:1018–21.

(20) Wei Q, Matanoski GM, Farmer ER, Hedayati M, Grossman L. DNA repair and aging in basal cell carcinoma: a molecular epidemiology study. Proc Natl Acad Sci U S A 1993;90:1614–8.

(21) Wiencke JK, Kelsey KT, Varkonyi A, Semey K, Wain JC, Mark E, et al.

Correlation of DNA adducts in blood mononuclear cells with tobacco carcinogen-induced damage in human lung. Cancer Res 1995;55:4910–4.
(22) Wei Q, Cheng L, Hong WK, Spitz MR. Reduced DNA repair capacity in lung cancer patients. Cancer Res 1996;56:4103–7.
(23) Hudmon KS, Honn SE, Jiang H, Chamberlain RM, Xiang W, Ferry G, et al. Identifying and recruiting healthy control subjects from a managed care organization: a methodology for molecular epidemiological case–control studies of cancer. Cancer Epidemiol Biomarkers Prev 1997;6:565–71.
(24) Athas AF, Hedayati M, Matanoski GM, Farmer ER, Grossman L. Development and field-test validation of an assay for DNA repair in circulating human lymphocytes. Cancer Res 1991;51:5786–93.
(25) Foecking MK, Hofstetter H. Powerful and versatile enhancer–promoter unit for mammalian expression vectors. Gene 1986;45:101–5.
(26) Moshell AN, Tarone RE, Newfield SA, Andrews AD, Robbins JH. A simple and rapid method for evaluating the survival of xeroderma pigmentosum lymphoid lines after irradiation with ultraviolet light. In Vitro 1981; 17:299–307.
(27) Barret JM, Calsou P, Salles B. Deficient nucleotide excision repair activity in protein extracts from normal human lymphocytes. Carcinogenesis 1995; 16:1611–6.
(28) Shivji MK, Kenny MK, Wood RD. Proliferating cell nuclear antigen is required for DNA excision repair. Cell 1992;69:367–74.
(29) McCuchan JH, Pagano JS. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. J Natl Cancer Inst 1968;41:351–7.
(30) Koch KS, Fletcher RG, Grond MP, Inyang AI, Lu XP, Brenner DA, et al. Inactivation of plasmid reporter gene expression by one benzo(*a*)pyrene diol-epoxide DNA adduct in adult rat hepatocytes. Cancer Res 1993;53(10 Suppl):2279–86.
(31) Sellers TA, Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. Increased familial risk for non-lung cancer among relatives of lung cancer patients. Am J Epidemiol 1987;126:237–46.
(32) Grossman L, Wei Q. DNA repair and epidemiology of basal cell carcinoma. Clin Chem 1995;41:1854–63.
(33) Loeb LA. Microsatellite instability: marker of a mutator phenotype in cancer. Cancer Res 1994;54:5059–63.
(34) Vulimiri SV, Wu X, Baer-Dubowska W, de Andrade M, Detry M, Spitz MR, et al. Analysis of aromatic DNA adducts and 7,8-dihydro-8-oxo-2′-deoxyguanosine in lymphocyte DNA from a case–control study of lung cancer involving minority populations. Mol Carcinog 2000;27:34–46.
(35) Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung cancer than male smokers? A case–control analysis by histologic type. Am J Epidemiol 1993;138: 281–93.
(36) Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. J Natl Cancer Inst 1996;88:183–92.
(37) O'Brien SN, Larcom LL, Baxley EG. Correlates of plasma cortisol and DNA repair in human peripheral lymphocytes: suppression of repair in women taking estrogen. Horm Res 1993;39:241–6.
(38) Ryberg D, Hewer A, Phillips DH, Haugen A. Different susceptibility to smoking-induced DNA damage among male and female lung cancer patients. Cancer Res 1994;54:5801–3.
(39) Tang DL, Rundle A, Warburton D, Santella RM, Tsai WY, Chiamprasert S, et al. Associations between both genetic and environmental biomarkers and lung cancer: evidence of a greater risk of lung cancer in women smokers. Carcinogenesis 1998;19:1949–53.
(40) Shriver SP, Bourdeau HA, Gubish CT, Tirpak DL, Davis AL, Luketich JD, et al. Sex-specific expression of gastrin-releasing peptide receptor: relationship to smoking history and risk of lung cancer. J Natl Cancer Inst 2000;92:24–33.
(41) Wei Q. Effect of aging on DNA repair and skin carcinogenesis: a minireview of population-based studies. J Investig Dermatol Symp Proc 1998;3: 19–22.
(42) Eller MS, Maeda T, Magnoni C, Atwal D, Gilchrest BA. Enhancement of DNA repair in human skin cells by thymidine dinucleotides: evidence for a p53-mediated mammalian SOS response. Proc Natl Acad Sci U S A 1997;94:12627–32.
(43) Cheng L, Eicher SA, Guo Z, Hong WK, Spitz MR, Wei Q. Reduced DNA repair capacity in head and neck cancer patients. Cancer Epidemiol Biomarkers Prev 1998;7:465–8.
(44) Wei Q, Eicher SA, Guan Y, Cheng L, Xu J, Young LN, et al. Reduced expression of hMLH1 and hGTBP/hMSH6: a risk factor for head and neck cancer. Cancer Epidemiol Biomarkers Prev 1998;7:309–14.
(45) Kovacs E, Almendral A. Reduced DNA repair synthesis in healthy women having first degree relatives with breast cancer. Eur J Cancer Clin Oncol 1987;23:1051–7.
(46) Pero RW, Johnson DB, Markowitz M, Doyle G, Lund-Pero M, Seidegard J, et al. DNA repair synthesis in individuals with and without a family history of cancer. Carcinogenesis 1989;10:693–7.
(47) Kolodner RD. Mismatch repair: mechanisms and relationship to cancer susceptibility. Trends Biochem Sci 1995;20:397–401.
(48) Yang LL, Maher VM, McCormick JJ. Relationship between excision repair and the cytotoxic and mutagenic effect of the 'anti' 7,8-diol-9,10-epoxide of benzo[*a*]pyrene in human cells. Mutat Res 1982;94:435–47.
(49) Cheng L, Bucana CD, Wei Q. Fluorescence *in situ* hybridization method for measuring transfection efficiency. Biotechniques 1996;21:486–91.
(50) Wang Y, Ichiba M, Iyadomi M, Zhang J, Tomokuni K. Effects of genetic polymorphism of metabolic enzymes, nutrition, and lifestyle factors on DNA adduct formation in lymphocytes. Ind Health 1998;36:337–46.
(51) Mooney LA, Bell DA, Santella RM, Van Bennekum AM, Ottman R, Paik M, et al. Contribution of genetic and nutritional factors to DNA damage in heavy smokers. Carcinogenesis 1997;18:503–9.
(52) Li D, Wang M, Cheng L, Spitz MR, Hittelman WN, Wei Q. *In vitro* induction of benzo(*a*)pyrene diol epoxide–DNA adducts in peripheral lymphocytes as a susceptibility marker for human lung cancer. Cancer Res 1996;56:3638–41.
(53) Wei Q, Gu J, Cheng L, Bondy ML, Jiang H, Hong WK, et al. Benzo(*a*) pyrene diol epoxide-induced chromosomal aberrations and risk of lung cancer. Cancer Res 1996;56:3975–9.
(54) Shen MR, Jones IM, Mohrenweiser H. Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Res 1998;58:604–8.

NOTES

Supported by Public Health Service grants R29CA70334 and R01CA74851 (to Q. Wei), R01CA55769 (to M. R. Spitz), and U19CA68437 (to W. K. Hong) from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. W. K. Hong is an American Cancer Society Clinical Research Professor.

We thank Lawrence Grossman (The Johns Hopkins University, Baltimore, MD) for providing pCMV*cat* and scientific advice, Maureen Goode (Department of Scientific Publications, The University of Texas M. D. Anderson Cancer Center, Houston) for editing the manuscript, Susan Honn for patient recruitment, Wayne Gosbee for data management, Yongli Guan for technical assistance, and Joanne Sider and Joyce Brown for assistance in preparing the manuscript.

Manuscript received March 30, 2000; revised August 25, 2000; accepted August 29, 2000.