Polymorphisms in the DNA Repair Genes XRCC1 and ERCC2, Smoking, and Lung Cancer Risk

Wei Zhou, Geoffrey Liu, David P. Miller, Sally W. Thurston, Li Lian Xu, John C. Wain, Thomas J. Lynch, Li Su, and David C. Christiani

Occupational Health Program, Departments of Environmental Health [W. Z., G. L., D. P. M., L. L. X., L. S., D. C. C.] and Biostatistics [S. W. T.], Harvard School of Public Health, Boston, Massachusetts, 02115, and Hematology-Oncology [G. L., T. J. L.], Pulmonary and Critical Care Unit [J. D. C. C.], Department of Medicine, and Thoracic Surgery Unit, Department of Surgery [J. C. W.], Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

Abstract

XRCC1 (X-ray cross-complementing group 1) and ERCC2 (excision repair cross-complementing group 2) are two major DNA repair proteins. Polymorphisms of these two genes have been associated with altered DNA repair capacity and cancer risk. We have described statistically significant interactions between the ERCC2 polymorphisms (Asp312Asn and Lys751Gln) and smoking in lung cancer risk. In this case-control study of 1091 Caucasian lung cancer patients and 1240 controls, we explored the gene-environment interactions between the XRCC1 Arg399Gln polymorphism, alone or in combination with the two ERCC2 polymorphisms, and cumulative smoking exposure in the development of lung cancer. The results were analyzed using logistic regression models, adjusting for relevant covariates. Overall, the adjusted odds ratio (OR) of XRCC1 Arg399Gln polymorphism (Gln/Gln versus Arg/Arg) was 1.3 (95% confidence interval (CI), 1.0–1.8). Stratified analyses revealed that the ORs decreased as pack-years increased. For nonsmokers, the adjusted OR was 2.4 (95% CI, 1.2–5.0), whereas for heavy smokers (≥55 pack-years), the OR decreased to 0.5 (95% CI, 0.3–1.0). When the three polymorphisms were evaluated together, the adjusted ORs of the extreme genotype combinations of variant alleles (individuals with 5 or 6 variant alleles) versus wild genotype (individuals with 0 variant alleles) were 5.2 (95% CI, 1.7–16.6) for nonsmokers and 0.3 (95% CI, 0.1–0.8) for heavy smokers, respectively. Similar gene-smoking interaction associations were found when pack-years of smoking (or smoking duration and smoking intensity) was fitted as a continuous variable. In conclusion, cumulative cigarette smoking plays an important role in altering the direction and magnitude of the associations between the XRCC1 and ERCC2 polymorphisms and lung cancer risk.

Introduction

Tobacco smoke contains an array of potent chemical carcinogens and reactive oxygen species that may produce DNA bulky adducts, cross-links, oxidative or base DNA damage, and DNA strand breaks. Among the several major DNA repair pathways that operate on specific types of damaged DNA by cigarette smoking, base-excision repair is involved in repair of DNA base damage and single strand breaks, and NER is involved in the repair of bulky monoaducts, cross-links, and oxidative damages (1-3).

XRCC1, 1 of >20 genes that participate in the base-excision repair pathway, has multiple roles in repairing DNA base damage and single-strand DNA breaks (4). Inconsistent results have been reported regarding the associations between the Arg399Gln (exon 10) polymorphism of XRCC1 and either functional significance or the risk of tobacco associated cancers (5-16). The Gln allele of this polymorphism was associated with higher levels of DNA adducts (5, 11) and glycophorin A variants (5), increased sister chromatid exchange frequencies (6, 7), and higher sensitivity to ionizing radiation (8); however, two other studies found no association between this polymorphism and elevated DNA adduct levels (9, 10). ERCC2 (also known as XPD), 1 of >20 genes that participate in the NER pathway, is involved in transcription-coupled NER and in the removal of a variety of structurally unrelated DNA lesions (17). Contradictory results are also reported on whether the functional significance or lung cancer risk of the Asp312Asn (exon 10) and Lys751Gln (exon 23) polymorphisms of ERCC2 (7, 10, 11, 14, 18-23).

Cigarette smoking may induce DNA damage, and individuals with a reduced DRC have a high level of carcinogenic-DNA adducts in their tissues (24). Lung cancer patients may have lower DRC when compared with healthy subjects (19, 25). The profound role of cigarette smoking in both lung cancer development and DNA damage suggests that smoking may be more than a simple confounding variable. Recently, we described statistically significant gene-environment interactions between the Asp312Asn or Lys751Gln polymorphism of ERCC2 and cumulative cigarette smoking in lung cancer risk, arguing for a possible biological interaction (26). Several other smaller studies found gene-smoking interactions for either XRCC1 or

3 The abbreviations used are: NER, nucleotide excision repair; XRCC1, X-ray cross-complementing group 1; ERCC2, excision repair cross-complementing group 2; DRC, DNA repair capacity; OR, odds ratio; CI, confidence interval.
ERCC2 polymorphisms in lung cancer risk (14–16, 23). Thus, the primary aim of this study was to determine whether cumulative smoking exposure altered the relationship between the XRCCI Arg399Gln polymorphism and lung cancer risk in a gene-environment interaction analysis. If XRCCI genotypes were found to interact with cigarette smoking exposure, a secondary aim of this study was to determine the joint association between XRCCI (Arg399Gln) and ERCC2 (Asp312Asn and Lys751Gln) polymorphisms and lung cancer risk, after accounting for potential interactions with cigarette smoking.

Materials and Methods

Study Population. The study was approved by the Human Subjects Committees of Massachusetts General Hospital and the Harvard School of Public Health. Details of this case-control population have been described previously (26–30). In brief, all of the histologically confirmed, newly diagnosed lung cancer patients at Massachusetts General Hospital were recruited between December 1992 and December 2000. Controls were recruited among friends and nonblood-related family members of the cases (usually spouses), with no specific matching characteristics. In some cases, if friends of lung cancer patients were not available, controls were recruited from friends and family of patients receiving thoracic surgery, chemotherapy, or radiation treatment for a condition other than lung cancer. Interviewer-administered questionnaires collected information on demographic and detailed smoking histories from each subject.

XRCCI and ERCC2 Genotyping. DNA was extracted from peripheral blood samples using the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN). The XRCCI Arg399Gln polymorphism was detected using modified PCR-restriction fragment length polymorphism methods, and published primer sequences (7). In brief, a 242-bp PCR product that included the Arg/Gln (A→C) allele in exon 10 (codon 399) was amplified, followed by MspI enzyme digestion (New England BioLabs, Beverly, MA).

The genotyping methods for the ERCC2 Asp312Asp (exon 10) and Lys751Gln (exon 23) polymorphisms have been described in detail (7, 26). Briefly, two separate PCR assays were used to detect the polymorphisms in exon 10 and exon 23 of ERCC2 using published primer sequences. DpnII and MspI (for exon 10), and Mbo II enzyme (for exon 23) digestion (New England BioLabs) were used for restriction fragment analyses. For quality control, a random 5% of the samples were repeated to assess the reproducibility of results. Two authors independently reviewed 100% of the agarose gels, with a third author arbitrating inconsistencies.

Statistical Analysis. Caucasian are 96% of the current samples. We analyzed all of the Caucasians with complete information on age, gender, smoking status (non-, ex-, and current smokers), pack-years of smoking, and years since smoking cessation (if ex-smoker). Years since smoking cessation were defined as zero for both current and nonsmokers.

In the gene-smoking interaction analysis, we used multiple approaches to evaluate consistency of results, including crude and adjusted analyses in specific categories of cumulative smoking exposure (i.e., pack-years), and genotype-smoking interactions models that considered separate continuous variables for cumulative smoking exposure (i.e., square root of pack-years, smoking duration, and log-transformed cigarettes per day). We fit the interactions between the XRCCI polymorphism and square root of pack-years, or various genotype combinations of XRCCI and ERCC2 polymorphisms and square root of pack-years in separate gene-environment interaction models. The interaction between smoking status and square root of pack-years was also included in all of the gene-smoking interaction models when using the square root of pack-years as a continuous variable, because it was found to be statistically significant in previous analyses of this population (26, 28, 29). Where appropriate, OR and 95% CI for the risk of lung cancer were calculated from these models. Statistical analyses were all undertaken using the SAS statistical packages (SAS Institute, Cary, NC).

Results

Baseline Characteristics. There were no significant demographic differences (age and gender) between enrolled and unenrolled eligible cases (>87% participation rate) and controls (>90% participation rate). A total of 2574 (99.2%) of 2597 enrolled subjects were genotyped successfully for all of the XRCCI (Arg399Gln) and ERCC2 (Asp312Asn and Lys751Gln) polymorphisms. We restricted our analysis to the 2331 Caucasians with complete data on age, gender, and smoking variables. Of these, there were 1091 lung cancer cases and 1240 controls. There was 100% concordance of randomly repeated samples and 99.7% agreement in independent gel interpretation between two individuals.

Detailed demographic characteristics including cigarette smoking history of cases and controls have been reported previously (26). Compared with control subjects, cases were older, more likely to be male, more likely to be current smokers, accumulated significantly greater pack-years of smoking, and had fewer years since smoking cessation. The mean age (and SD) of all of the cases was 64.9 ± 10.8 years (range, 26–91), and of controls, 58.5 ± 12.4 years (range, 19–100). Females represented 46.4% of cases and 54.4% of controls. Among cases, the mean pack-years of ex-smokers was 55.3 ± 36.0, and of current smokers, 64.2 ± 35.8; among controls, the numbers were 29.3 ± 27.8 and 38.1 ± 25.2 respectively. The years since smoking cessation for ex-smokers were 14.2 ± 11.0 and 19.4 ± 11.9 for cases and controls, respectively.

The distribution of smoking variables in our controls was similar to the general Massachusetts population over age 45. The proportions of non-, ex-, and current smokers were 35.0, 45.5, and 19.5% in our controls, and 36.0, 47.0, and 17.0% in the general Massachusetts population over age 45 years, respectively. For current smokers, mean cigarettes per day (controls: 21.2 cigarettes; Massachusetts: 21.4 cigarettes) and earliest age of smoking (controls: 17.9 years; Massachusetts: 17.9) were similar. For ex-smokers, the proportions of those who...
have quit smoking for >5 years were 87.4% (controls) and 85.5% (Massachusetts).

**Distribution of DNA Repair Gene Polymorphisms Among Cases and Controls.** The XRCC1 polymorphism in this control population was consistent with Hardy-Weinberg equilibrium ($P = 0.95$, $\chi^2$ goodness of fit). The frequencies of Arg/Arg, Arg/Gln, and Gln/Gln were 42.8, 42.9, and 14.3%, respectively, in cases, and 44.5, 44.0, and 11.5% in controls. For older subjects (age $\geq$55 years), the genotype frequencies were very similar between the 886 cases and the 769 controls, with the frequencies of 42.2, 44.1, and 13.7% in cases, and 42.8, 44.7, and 12.5% in controls, respectively. For younger subjects (age <55 years), cases have higher frequency of the Gln/Gln genotype than controls; the frequencies of Arg/Arg, Arg/Gln, and Gln/Gln were 45.4, 37.6, and 17.1%, respectively, in the 205 cases, and 47.1, 42.9, and 10.0% in the 471 controls, respectively. In categories of higher pack-years of smoking, the frequency of Arg/Arg genotype showed a trend of increase in cases ($P_{\text{trend}} = 0.06$ for general linear models) but not in controls ($P_{\text{trend}} = 0.43$; Table 1; two-step test; Refs. 26, 29, 33).

The frequency distributions of the ERCC2 Asp312Asn and Lys735Gln polymorphisms have been described previously; the controls are also in Hardy-Weinberg equilibrium (26). Trends in different pack-year categories of smoking for these two polymorphisms were similar to the XRCC1 polymorphism.

**Associations between the XRCC1 Arg399Gln Polymorphism and Lung Cancer Risk.** There was a borderline overall relationship between XRCC1 polymorphism and lung cancer risk. Compared with the Arg/Arg genotype, the adjusted ORs of lung cancer were 1.0 (95% CI, 0.8–1.3) for the Arg/Gln genotype, and 1.3 (95% CI, 1.0–1.8) for the Gln/Gln genotype, which were similar to the crude results (Table 1). Similar associations between this polymorphism and lung cancer risk were found for males and females separately, and for different histological (adenocarcinoma and squamous cell carcinoma) and clinical stages (stage I/II and stage IIIA/IIIB/IV) subgroups.

We found a statistically significant interaction ($P = 0.01$) between XRCC1 genotype (Gln/Gln versus Arg/Arg) and age subgroup (older versus younger) in the risk of lung cancer. The adjusted ORs of the Gln/Gln genotype versus the Arg/Arg genotype were 2.4 (95% CI, 1.3–4.2) for younger subjects (age $<55$ years; Ref. 30) and 1.0 (95% CI, 0.7–1.5) for older subjects (age $\geq55$ years). Similar trends were obtained when the age subgroups were dichotomized at 50, 60, or 65 years.

**Relationship among Cumulative Cigarette Smoking, DNA Repair Gene Polymorphisms, and Lung Cancer Risk.** For the XRCC1 polymorphism, compared with the Arg/Arg genotype, the Gln/Gln genotype was a risk factor for lung cancer in nonsmokers and light smokers (≤25 pack-years), but a protective factor in heavy smokers (>55 pack-years), both in crude and adjusted analyses (Table 1). Similar associations were found when we stratified the population into corresponding levels by tertiles or quartiles of pack-years.

For the gene-smoking interaction analysis between the XRCC1 polymorphism and lung cancer risk, the magnitude and statistical significance of the interaction term between genotype group and square root of pack-years for the Gln/Gln versus Arg/Arg genotype comparison was stronger, respectively, than that for the Arg/Gln versus Arg/Arg comparison, suggesting a dose-response relationship for the number of Gln alleles (Fig. 1).

Consistency of the results was evaluated by considering smoking variables in different ways. We decomposed pack-years into its component parts of smoking intensity (mean number of cigarettes per day) and duration (in years). Similar gene-smoking interaction associations were found when either

<table>
<thead>
<tr>
<th>n (case/control)</th>
<th>Percentage (case/control)</th>
<th>Crude ORs$^b$</th>
<th>Adjusted ORs$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arg/Arg</td>
<td>Arg/Gln</td>
<td>Gln/Gln</td>
</tr>
<tr>
<td>Total</td>
<td>1091/1240</td>
<td>42.8/44.4</td>
<td>42.9/44.0</td>
</tr>
<tr>
<td>Non-</td>
<td>73/434</td>
<td>35.6/44.0</td>
<td>43.8/45.4</td>
</tr>
<tr>
<td>Mild-</td>
<td>149/389</td>
<td>39.6/45.5</td>
<td>46.3/42.2</td>
</tr>
<tr>
<td>Moderate-</td>
<td>378/286</td>
<td>40.0/46.9</td>
<td>44.2/43.4</td>
</tr>
<tr>
<td>Heavy-</td>
<td>491/131</td>
<td>47.1/57.4</td>
<td>40.7/46.6</td>
</tr>
</tbody>
</table>

$^a$ Mild, moderate, and heavy smokers correspond to the three tertiles of pack-years in ever smokers for all participants. The tertiles were divided at 25 and 55 pack-years.

$^b$ Stratified analysis. In all logistic regression models, the Arg/Arg genotype was the reference group. For the adjusted analysis, the logistic regression model included the following covariates: age, gender, square root of pack-years, smoking status, time since smoking cessation (in years), and genotype groups.
log-transformed cigarettes per day or smoking duration (both as continuous variables) were substituted for square root of pack-years in the regression models. The gene-smoking interaction term between XRCC1 genotype (Gln/Gln versus Arg/Arg) and log-transformed cigarettes per day was statistically significant ($P = 0.03$), and the interaction between genotype and years of smoking (Gln/Gln versus Arg/Arg) was borderline significant ($P = 0.07$).

In a recent publication, we described similar genotype-smoking interaction associations for the ERCC2 Asp312Asn and Lys751Gln polymorphisms where there was a dose-response relationship for the number of variant Asn or Gln alleles (26). The variant Asn or Gln alleles were risk factors in nonsmokers, but protective factors in heavy smokers when compared with the Asp/Asp or Lys/Lys wild-type.

**Combined Effects of Two DNA Repair Genes on Lung Cancer Risk.** We dichotomized the population into six genotype groups based on the number of variant alleles of the three XRCC1 and ERCC2 polymorphisms: group 0, the reference group, with no variant alleles of either gene, *i.e.*, individuals with wild genotype for all of the three polymorphisms; group 1, with 1 variant allele; group 2, with 2 variant alleles; and so forth. The frequency distributions of each genotype combination group by different pack-year categories are shown in Table 2.

In all of the logistic regression analyses, the genotype group 5 and group 6 were combined into one group (group 5/6) because of the small sample sizes. For subjects of different pack-year categories of smoking, when compared with the reference genotype group (group 0), the genotype combination groups with $>$3 variant alleles were risk factors for lung cancer in nonsmokers but protective factors in heavy smokers (Table 3). In stratified regression models, the adjusted ORs of the extreme genotype combinations of variant alleles (group 5/6 versus group 0) were 5.2 (95% CI, 1.7–16.6) for nonsmokers, 2.2 (95% CI, 0.9–5.2) for mild smokers, 2.0 (95% CI, 0.8–5.1) for moderate smokers, and 0.3 (95% CI, 0.1–0.8) for heavy smokers, respectively. Table 3 presented the results of the combined effects of cigarette smoking and different genotype groups (XRCC1 Arg399Gln polymorphism alone or the joint XRCC1 and ERCC2 polymorphisms) on the risk of lung cancer.

For the formal gene-smoking interaction analysis between the XRCC1 and ERCC2 combined genotypes and lung cancer risk, the magnitude and statistical significance of the interaction

---

**Table 2** Genotype frequencies of the combined XRCC1 Arg399Gln, ERCC2 Asp312Asn, and ERCC2 Lys751Gln polymorphisms for different pack-years of smoking$^{a,b}$

<table>
<thead>
<tr>
<th>Variant alleles</th>
<th>n (case/control)</th>
<th>Percentage (case/control)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 0</td>
<td>Group 1</td>
</tr>
<tr>
<td>Total</td>
<td>109/1240</td>
<td>14/1.14.4</td>
</tr>
<tr>
<td>Non-</td>
<td>73/134</td>
<td>8/2.13/8</td>
</tr>
<tr>
<td>Mild-</td>
<td>149/399</td>
<td>16/1/72</td>
</tr>
<tr>
<td>Moderate-</td>
<td>378/286</td>
<td>13/5/14.3</td>
</tr>
<tr>
<td>Heavy-</td>
<td>491/131</td>
<td>14/9/7.6</td>
</tr>
</tbody>
</table>

$^a$ Wild, moderate, and heavy smokers correspond to the three tertiles of ever smokers, in ever smokers only.

$^b$ Combined genotype group 0 were the individuals with 0 variant alleles of the three polymorphisms (wild-type for all of the three polymorphisms); group 1 were the individuals with 1 variant allele; group 2 were the individuals with 2 variant alleles; group 3 were the individuals with 3 variant alleles; group 4 were the individuals with 4 variant alleles; group 5 were the individuals with 5 variant alleles; and group 6 were the individuals with 6 variant alleles.

---

**Table 3** Adjusted ORs (95% CI) for the joint effect of XRCC1 polymorphism (or combined XRCC1 and ERCC2 polymorphisms) and different smoking categories$^c$

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Non-Marginal Smoking</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>XRCC1 Arg399Gln</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg/Arg</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Arg/Gln</td>
<td>1.2 (0.7–2.1)</td>
<td></td>
</tr>
<tr>
<td>Gln/Gln</td>
<td>2.6 (1.2–5.4)</td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>XRCC1 Arg399Gln + ERCC2 Asp312Asn + ERCC2 Lys751Gln$^d$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>1.5 (0.5–4.3)</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>0.9 (0.3–2.7)</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>1.8 (0.6–4.8)</td>
<td></td>
</tr>
<tr>
<td>Group 4</td>
<td>3.8 (1.3–11.0)</td>
<td></td>
</tr>
<tr>
<td>Group 5/6</td>
<td>5.4 (1.7–17.4)</td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

$^c$ Interaction Models: separate logistic regression model for the XRCC1 Arg399Gln polymorphism alone or the combined three polymorphisms including the following covariates: age, gender, three dummy variables for pack-years, smoking status, time since smoking cessation (in years), genotype groups, interaction terms between genotype groups, and the three dummy variables of smoking. Logistic regression models for marginal calculations included the above covariates, except for the interaction terms. The ORs in different smoking categories are for current smokers or nonsmokers only. For ex-smokers, the ORs for each comparison are $~25\%$ less than for current smokers.

$^d$ Combined genotype group 0 were the individuals with 0 variant alleles of the three polymorphisms (wild-type for all of the three polymorphisms); group 1 were the individuals with 1 variant allele; group 2 were the individuals with 2 variant alleles; group 3 were the individuals with 3 variant alleles; group 4 were the individuals with 4 variant alleles; group 5 were the individuals with 5 variant alleles; and group 6 were the individuals with 6 variant alleles.
term between genotype group and square root of pack-years was strongest when comparing group 5/6 to group 0 (Fig. 2). The adjusted ORs of all five of the comparisons decreased as pack-years of smoking increased, and the fitted ORs were comparable with the results when pack-years was stratified into different categories. Similar gene-smoking interaction associations were found when either log-transformed cigarettes per day or smoking duration (both as continuous variables) were substituted for square root of pack-years in the regression models. The gene-smoking interaction terms between the extreme genotype combinations (group 5/6 vs. group 0) were the individuals with 5 or 6 variant alleles; group 4 were the individuals with 4 variant alleles; group 5/6 were the individuals with 5 or 6 variant alleles. The fitted ORs (——) and 95% CIs (—) represent the lung cancer risk of combined genotype group 1 (first on the left), group 2 (second on the left), group 3 (middle), group 4 (second on the right), or group 5/6 (first on the right), when each is compared with the combined genotype group 0, at different numbers of pack-years (PY). The logistic regression model included the following covariates: age, gender, square root of pack-years, smoking status, time since smoking cessation (in years), combined genotype groups, interaction terms between different combined genotype groups and square root of pack-years, and interaction between smoking status and square root of pack-years. The Ps for interaction terms between combined genotype groups and square root of pack-years were 0.13 for group 1, 0.61 for group 2, 0.02 for group 3, <0.01 for group 4, and <0.01 for group 5/6.

Discussion

In this study, we used cumulative cigarette smoking as a surrogate for tobacco carcinogen exposure. We found statistically significant interactions between the three polymorphisms of the two DNA repair genes (XRCC1 and ERCC2), both separately and combined, and cumulative cigarette smoking in lung cancer risk in this Caucasian sample. Invariably, the variant alleles of the three DNA repair genetic polymorphisms were risk factors in non- or light smokers, but tended to be protective in heavy smokers. This finding was consistent regardless of the analytical techniques and modeling for different subgroups by age, gender, histological cell types, and clinical stages.

In smaller studies, similar gene-smoking relationships have been described for the polymorphisms of both XRCC1 (Arg399Gln) and ERCC2 (Asp312Asn and Lys751Gln; Refs. 11, 14–16, 23). For XRCC1, one study with 308 healthy Italian individuals found that the Gln/Gln genotype had significantly higher DNA adduct levels in lymphocytes only in never-smokers but not in ever-smokers (11). Recent epidemiological studies involving Caucasian (180 cases), African-American (154 cases), or Korean (192 cases) suggested consistently that the variant Gln allele was associated with higher risk of lung cancer for lighter smokers and lower risk in heavier smokers (15, 16). For ERCC2, the Asp/Asp genotype of the Asp312Asn polymorphism was associated with higher lung cancer risk for light smokers only, and not for non- or heavier smokers in one study with 96 Caucasian cases (14). A recent study with 185 Swedish lung cancer cases suggested that the variant alleles of the two ERCC2 polymorphisms were associated with increased risk for never-smokers only, and tended to be protective in ever-smokers (23). In addition, inconsistent results were demonstrated regarding the possible functional consequences of polymorphisms of both DNA repair genes, and between individual polymorphisms and lung cancer risk (5–10, 12, 13, 18–22). Discrepancies across these studies may be because of relatively small sample sizes analyzed in the various subgroups or because of associations that were specific to different cumulative smoking levels.

This is the first report showing the gene-smoking interactions between the joint effects of XRCC1 and ERCC2 genotypes and cumulative cigarette smoking exposure in lung cancer risk. Our results suggest that the magnitudes of these gene-smoking interactions are associated with the number of variant alleles of these three polymorphisms (Table 3; Fig. 2). Comparison of the
extreme joint genotype combinations of XRCC1 and ERCC2 is consistent with an additive effect of the two polymorphic genes at each level of smoking exposure (data not shown).

The exact mechanism of how cigarette smoking changes the DRC posed by each genotype of these DNA repair genetic polymorphisms is unknown. One possible explanation may be that different DRCs of different genotypes are overwhelmed by heavy smoking exposure. Alternatively, cigarette smoking may stimulate DRC in response to the DNA damage caused by tobacco carcinogens, because heavy smokers among both lung cancer patients and controls may have more proficient DRC in lymphocytes than non- or light smokers (34), and DNA repair gene expression was increased in heavy smokers among both head and neck cancer patients and controls (35). Because heavy smokers tended to be older subjects (the frequencies of older subjects in non-, mild, moderate, and heavy smokers were 60.6, 60.4, 72.9, and 86.7%, respectively), the lower risk for the variant alleles in heavy smokers may reflect a prolonged survival for the lung cancer patients. A recent study has shown that effective host DRC may be associated with poorer survival in patients with non-small cell lung cancer who are treated with chemotherapy (36). We also found that the Glo/Gln genotype of the XRCC1 Arg399Gln polymorphism was associated with a markedly increased risk of lung cancer in younger individuals but not in older individuals. Additional analysis is needed to explore the effect of DNA repair genetic polymorphisms on the survival of lung cancer patients.

This study has a number of limitations. Firstly, our study was a hospital-based case-control study. Secondly, our controls were healthy spouses or friends of individuals with lung cancer or other cardiothoracic conditions. Because spouse or friend controls are more similar to cases than population controls because they are likely to share similar health behaviors, there may be some potential bias for our results. However, bias in the estimate of the stratum-specific OR because of a specific gene with strata defined by levels of measured confounders and effect modifiers (such as smoking, diet, and so forth) will only occur if, within these strata, spouse and friend controls are more likely than other types of controls to have a particular allele or alleles of the genotype under study. Although we did not match individually our controls to cases on age, gender, race, or smoking variables, we did adjust for these variables in the analyses and also performed stratified analyses by these variables, and each time found consistent gene-smoking associations. Thirdly, the exact biological mechanisms for the gene-smoking interaction associations are unclear, because there is no direct phenotype data for the function of these polymorphisms. Nonetheless the current study adds weight to existing epidemiological studies that have found a similar gene-smoking interaction association between these two DNA repair genes and cumulative cigarette smoking (11, 14–16, 23). Fourthly, we only evaluated the Arg399Gln polymorphism of XRCC1, and Asp312Asn and Lys751Gln polymorphisms of ERCC2 in this study, and did not evaluate the other polymorphisms of these two genes, including the XRCC1 Arg194Thr polymorphism, because of scarce functional data and relatively low allele frequencies for these polymorphisms (5, 6, 14, 15, 35, 36), which may result in some misclassification in the functions of these two DNA repair genes. Besides, XRCC1 and ERCC2 only contribute partially to DNA repair capacity in their respective pathways, and it is possible that polymorphisms of other genes not evaluated in this study could play a role in determining lung cancer risk. However, evaluation of a greater number of polymorphisms will require large sample sizes. Fifthly, the data of environmental tobacco smoking and alcohol assumption are not adjusted in our analysis. In the case of environmental tobacco smoking, the use of spouse and friend controls may actually mitigate such differences, given that the environment of cases and controls may be generally similar. Alcohol assumption may modify the risk of head and neck, or esophageal cancers, whereas a recent review suggests that the evidence is far from conclusive in the risk of lung cancer (37). The above limitations may be one reason for the inconsistent trend in risk of lung cancer by XRCC1 alone or the combined XRCC1 and ERCC2 genotypes across different smoking levels; i.e., the OR of XRCC1 Gln/Gln versus Arg/Arg in moderate smoking group is bigger than that in the light smoking group (Table 1).

In conclusion, cumulative cigarette smoking exposure appears to play an important role in altering the direction and magnitude of the association between XRCC1 and ERCC2 polymorphisms, and the risk of lung cancer. Lung cancer risk by the polymorphisms of both XRCC1 and ERCC2 is dependent on cumulative smoking exposure. Additional studies are needed to detect the function of these polymorphisms, and its associations with cigarette smoking and clinical prognosis.

Acknowledgments
We thank the following staff members of the Lung Cancer Susceptibility Group: Barbara Bean, Jessica Shen, Andrea Solomon, Andrea Shaffer, So-Aye Shaneyfelt, Lucy Ann Principe, Salvatore Mucci, Richard Rivera-Massa, Lisa I. Wang, Stephanie Shih, and Maria Fragosso; and Dr. Panos Fidias, and the physicians and surgeons of the Massachusetts General Hospital Cancer Center for generous support.

References


Polymorphisms in the DNA Repair Genes XRCC1 and ERCC2, Smoking, and Lung Cancer Risk

Wei Zhou, Geoffrey Liu, David P. Miller, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/12/4/359

Cited articles
This article cites 35 articles, 14 of which you can access for free at:
http://cebp.aacrjournals.org/content/12/4/359.full#ref-list-1

Citing articles
This article has been cited by 22 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/12/4/359.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cebp.aacrjournals.org/content/12/4/359.
Click on "Request Permissions" which will take you to the Copyright Clearance Center’s (CCC) Rightslink site.