Reduced expression levels of nucleotide excision repair genes in lung cancer: a case-control analysis

Lie Cheng1,2, Margaret R.Spitz1, Waun K.Hong2 and Qingyi Wei1,4

1Department of Epidemiology and 2Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

3Present address: Institute of Molecular Biology, The University of Texas Center of Human Genetics, Houston, TX 77030, USA

4To whom correspondence should be addressed
Email: qwei@notes.mdacc.tmc.edu

Epidemiological studies have indicated that reduced DNA repair capacity and increased DNA adduct levels are associated with increased risk of lung cancer. Nucleotide excision repair (NER) is the major pathway in humans for repairing DNA adducts induced by smoking-related carcinogens, such as benzo[a]pyrene diolepoxide. We hypothesized that genetically determined baseline expression level of genes involved in NER is associated with risk of lung cancer. In a pilot case-control study, we measured the relative expression levels of five NER genes [ERCC1, XPB/ERCC3, XPG/ERCC5, CSB/ERCC6 and XPC (ERCC, excision repair cross-complementing; CSB, Cockayne’s syndrome complementary group B)] in phytohemagglutinin-stimulated peripheral lymphocytes obtained from 75 lung cancer patients and 95 controls using a newly developed multiplex RT–PCR assay. Cases and controls were matched on age, sex, ethnicity and tobacco use. The expression level of the β-actin gene was used as an internal control for the relative quantitation. We observed a 12.2 and 12.5% decrease in the baseline expression levels of XPG/ERCC5 and CSB/ERCC6, respectively, in cases compared with controls. These differences were statistically significant (P < 0.01) when the median expression level in the controls was used as the cut-off point, the lung cancer patients were significantly more likely than the controls to have reduced expression levels of XPG/ERCC5 [odds ratio (OR), 2.32; 95% confidence interval (CI), 1.22–4.43] and CSB/ERCC6 (OR, 2.49; 95% CI, 1.28–4.84). There was also a dose–response relationship between reduced expression levels and increased lung cancer risk (trend test: P < 0.01). Our results suggest that individuals whose expression levels of XPG/ERCC5 and CSB/ERCC6 are reduced may be at higher risk of lung cancer.

Introduction

Smoking is the major cause of cancer deaths in man. The relative risk of lung cancer in current smokers is 20-fold greater than in non-smokers (1). Because tobacco carcinogens cause DNA damage, reduced DNA repair capacity plays a role in the etiology of tobacco-related cancers (2,3). One major DNA repair pathway capable of removing a variety of structurally unrelated DNA lesions, including those induced by tobacco carcinogens, is nucleotide excision repair (NER) (4). This complex DNA repair process consists of approximately 30 proteins involved in sequential damage recognition, chromatin remodeling, incision of the damaged DNA strand on both sides of the lesion, excision of the oligonucleotide containing the damage and gap-filling DNA synthesis followed by strand ligation (5). For example, smoking-related bulky adducts induced by benzo[a]pyrene or other polyaromatic hydrocarbons and arylamines are removed effectively by the NER pathway (6). In xeroderma pigmentosum (XP), patients have an extraordinarily higher rate of skin cancer because of a genetically determined defect in NER (4). Other cancer-prone patients who have deficient DNA repair also have a higher rate of internal cancer (7).

Physiologically, the DNA repair capacity should be correlated with the level of proteins involved in DNA repair activity, which is controlled at the transcriptional level (4). Therefore, it is conceivable that the baseline transcript level of DNA repair genes reflects a cellular ability to meet repair demand once the cells are stimulated by carcinogen exposure. Although reduced DNA repair capacity has been shown to be associated with smoking-related cancers such as lung cancer (2) and head and neck cancer (3), few studies have evaluated the role of expression of DNA repair genes in the etiology of lung cancer. We hypothesized that the genetically determined baseline expression level of genes involved in NER is associated with risk of lung cancer. To test this hypothesis and use a newly developed multiplex RT–PCR assay (8), we conducted a molecular epidemiological study in which we measured the relative expression levels of five NER genes [ERCC1, XPB/ERCC3, XPG/ERCC5, CSB/ERCC6 and XPC (ERCC, excision repair cross-complementing; CSB, Cockayne’s syndrome complementary group B)] in phytohemagglutinin-stimulated peripheral lymphocytes obtained from 75 patients having newly diagnosed, untreated lung cancer and 95 cancer-free controls.

Materials and methods

Study subjects

Lung cancer patients and controls were selected from an ongoing lung cancer study at The University of Texas M.D. Anderson Cancer Center. The 75 patients had histopathologically confirmed lung cancer and had not received prior chemotherapy or radiotherapy. The 95 controls were a subset of comparison subjects selected from the clinics of a large, multispecialty health maintenance organization (9) and matched with the lung cancer patients on age (±5 years), sex, ethnicity and smoking status. After informed consent was obtained, each subject was interviewed using a standard questionnaire which requested information on age, sex, ethnicity and tobacco use. In addition, each subject donated 20 ml of blood in heparinized tubes. The research protocol was approved by the M.D. Anderson Cancer Center Institutional Review Board.

Blood sample processing and culturing

Lymphocyte cultures were prepared as described previously (10). Briefly, 1 ml of fresh whole blood obtained from each subject was inoculated into one T-25 flask containing 9 ml of RPMI-1640 medium supplemented with 20% fetal

Abbreviations: CI, confidence interval; CSB, Cockayne’s syndrome complementary group B; ERCC, excision repair cross-complementing; NER, nucleotide excision repair; OR, odds ratio; PHA, phytohemagglutinin.
bovine serum and 112.5 µg/ml PHA and then incubated at 37°C for 72 h. Cell culture should reduce any transient effect of serum factors on the gene expression. Stimulation by the mitogen, PHA, is also necessary (8), because current smokers, including recent quitters (within 1 year previously). Adjusted the temperatures as well as adjusting the concentrations of each of the primers (RNA were collected by Ficoll gradient centrifugation and washed with phosphate-buffered saline; the cell pellet was then suspended in 1 ml Tri-Reagent™ (RNA–DNA/Protein Isolation Reagent; Molecular Research Center, Cincinnati, OH) and stored at −80°C for later RNA extraction.

**Multiplex RT–PCR**

The multiplex RT–PCR was used to amplify the five NER genes as described previously (8). Briefly, total RNA was extracted from PHA-stimulated lymphocytes using Tri-Reagent™. For synthesis of cDNA, the isolated mRNA was reverse-transcribed using random primers (Promega Biotech, Piscataway, NJ) and Moloney murine leukemia virus RT (MMLV RT; United States Biochemical Co., Cleveland, OH). A 20 µl RT reaction mixture was prepared for each sample and contained 2 µg total cellular RNA, 4 µl 5 × RT buffer (250 mM Tris–HCl pH 8.3, 375 mM KCl, 50 mM DTT, 15 mM MgCl₂) (Bethesda Research Laboratory, Gaithersburg, MD), 0.1 mM dNTP, 1.0 µg random primer, 20 U RNasin (Promega Biotech, Madison, WI), 100 U MMLV RT and 6.5 µl diethyl pyrocarbonate-treated water (Promega Biotech) that was free of RNase. The mixture was incubated at room temperature for 10 min, kept at 42°C for 45 min, heated to 95°C for 10 min, and then quick-chilled on ice.

PCR was optimized by selecting unique primer sequences and annealing temperatures as well as adjusting the concentrations of each of the primers (10). Briefly, PCR was performed using a total volume of 40 µl of a final concentration of 1× PCR buffer (500 mM KCl, 100 mM Tris–HCl pH 9.0, 1% Triton X-100, 2.5 mM MgCl₂), 0.1 mM each dNTP, 2 U Taq polymerase (Promega Biotech), 75 pmol β-actin primers, 35 pmol ERCC1 primers, 50 pmol XPB/ERCC3 primers, 50 pmol ERF/ERCC5 primers, 50 pmol CSB/ERCC6 primers and 25 pmol XPC primers. The reaction mixture was heated at 95°C for 5 min. The amplification profile involved an initial denaturing step at 95°C for 5 min, 29 cycles of denaturation at 95°C for 30 s, primer annealing at 59°C for 30 s and extension at 72°C for 45 s, and a final elongation step at 72°C for 10 min.

The PCR products were electrophoresed in a 2% agarose gel, stained with ethidium bromide to visualize the bands under UV irradiation and photographed using either regular Polaroid film (Polaroid Corp., Cambridge, MA) or a computerized imaging system (Model IS-1000; Alpha Innotech Co., San Leandro, CA). The bands on the negative image were scanned and analyzed by densitometry; the areas of the peaks were calculated in arbitrary units. To evaluate the relative levels of expression of the target genes coamplified in the same PCR, the value of the internal standard (β-actin) in each sample was used as the background measurement (1.00) of cellular gene expression in sample. The relative value was generated by calculating the ratio of the arbitrary units of each target gene to that of β-actin in the same PCR.

**Statistical analysis**

Student’s t-test was used to compare the differences in the relative expression levels analyzed as a continuous variable between groups. For calculation of crude ORs and CIs, the median relative expression level in the controls was used as the cut-off point for each gene. Ever smokers were defined as those who had smoked more than 100 cigarettes in a lifetime; they were further divided into former smokers who had quit smoking >1 year previously and current smokers, including recent quitters (within 1 year previously). Adjusted ORs were calculated by fitting logistical regression models with adjustment for age, sex, race and smoking status. All the statistical analyses were performed using SAS software (version 6; SAS Institute Inc., Cary, NC).

**Results**

The demographics of the subjects are given in Table I. The mean ages of the patients and controls were 58.6 ± 10.6 and 60.2 ± 10.3 years, respectively; the difference between the two means was not statistically significant. There was no significant difference in the distribution of sex, ethnicity and smoking status between cases and controls, suggesting that the matching was adequate.

Using serial dilutions of cDNA, we have demonstrated previously the linearity of multiplex RT–PCR amplification of the five NER genes and β-actin gene over 25–30 PCR cycles (10). Therefore, we selected 29 cycles for all PCR reactions in the present study. We were able to measure expression levels of all five NER genes in PHA-stimulated lymphocytes by multiplex RT–PCR, suggesting that all the five target genes with the β-actin gene in the same tube could be simultaneously coamplified without producing other non-specific products.

An example of the gene expression pattern in the patients and controls is shown in Figure 1. There was considerable difference in the distribution of sex, ethnicity and smoking status (data not shown). The differences in the mean relative expression levels of the five genes between lung cancer patients and healthy controls is shown in Table I. There was considerable variation in the relative levels of expression of the five selected genes. Using the β-actin expression level as a reference (baseline expression, 100%), the relative expression levels of the five genes were consistently lower in PHA-stimulated lymphocytes obtained from lung cancer patients than in those from controls (Figure 2). The differences in the mean relative expression levels of the five genes between lung cancer patients and controls are listed in Table II. Statistically significant reductions in the expression of XPG/ERCC5 (~12.2%, P = 0.001) and CSB/ERCC6 (~12.5%, P = 0.004) were found in lung cancer patients as compared with that in controls. There was no apparent association between gene expression and age, sex, ethnicity or smoking status (data not shown).

In logistic regression models including age, sex, ethnicity, smoking status and the relative gene expression levels, a significantly increased risk of lung cancer was associated with reduced expression of XPG/ERCC5, CSB/ERCC6 and XPC (P < 0.01). Individuals having low expression levels of these genes exhibited a 2-fold greater risk of lung cancer after adjustment for age, sex, ethnicity and smoking status (Table III).
regulated gene expression is DNA methylation status. For XPC/ERCC6, XPG/ERCC5, XPB/ERCC3, the reduced NER gene expression pattern in lung cancer patients and controls to increase concomitantly. The results of the present study are shown in Table II.

Table II. Differences in NER gene expression between lung cancer patients and controls

<table>
<thead>
<tr>
<th>Gene</th>
<th>Patients (n=75)</th>
<th>Controls (n=95)</th>
<th>Percent difference</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERCC1</td>
<td>53.9 ± 18.5</td>
<td>58.8 ± 18.9</td>
<td>-8.3</td>
<td>0.091</td>
</tr>
<tr>
<td>ERCC3</td>
<td>54.5 ± 15.2</td>
<td>57.9 ± 16.5</td>
<td>-5.8</td>
<td>0.172</td>
</tr>
<tr>
<td>ERCC5</td>
<td>60.3 ± 20.5</td>
<td>68.7 ± 19.0</td>
<td>-12.2</td>
<td>0.001</td>
</tr>
<tr>
<td>ERCC6</td>
<td>55.3 ± 16.6</td>
<td>63.2 ± 18.1</td>
<td>-12.5</td>
<td>0.004</td>
</tr>
<tr>
<td>XPC</td>
<td>50.4 ± 19.2</td>
<td>54.7 ± 20.0</td>
<td>-7.8</td>
<td>0.162</td>
</tr>
</tbody>
</table>

Data represent means ± SD.
Percent difference [(patient-control)/control]×100.
Two-sided t-test.

Table III. Crude and adjusted ORs and 95% CIs for NER gene expression levels in lung cancer patients and controls

<table>
<thead>
<tr>
<th>Gene expression level</th>
<th>Cases (n=75)</th>
<th>Controls (n=95)</th>
<th>Crude OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERCC1</td>
<td>29</td>
<td>48</td>
<td>1.00</td>
<td>1.76 (0.93–3.33)</td>
</tr>
<tr>
<td>XPB/ERCC3</td>
<td>30</td>
<td>48</td>
<td>1.00</td>
<td>1.56 (0.84–2.90)</td>
</tr>
<tr>
<td>XPG/ERCC5</td>
<td>23</td>
<td>47</td>
<td>1.00</td>
<td>2.32 (1.22–4.43)</td>
</tr>
<tr>
<td>CSB/ERCC6</td>
<td>21</td>
<td>47</td>
<td>1.00</td>
<td>2.49 (1.28–4.84)</td>
</tr>
<tr>
<td>XPC</td>
<td>23</td>
<td>47</td>
<td>1.00</td>
<td>2.19 (1.15–4.18)</td>
</tr>
</tbody>
</table>

*Adjusted in a logistic regression model including age in years, sex, ethnicity and smoking status.

Discussion

In this study, we measured the relative expression levels of five NER genes in PHA-stimulated peripheral blood lymphocytes obtained from 75 lung cancer patients and 95 controls using a newly developed multiplex RT-PCR assay. We found that lung cancer patients had significantly reduced expression levels of XPG/ERCC5 and CSB/ERCC6 when compared with those of the controls, suggesting that low expression level of these NER genes may be associated with an increased risk of lung cancer. These findings are biologically plausible. In addition, we have demonstrated previously that reduced DNA repair capacity as measured by the host-cell reactivation assay is associated with an increased risk of lung cancer (3). Based on these new findings, we postulate that the previously observed lower DNA repair capacity in lung cancer patients may be the consequence of low expression level of these NER genes observed in the present study. However, the direct relationship between DNA repair capacity and expression level of these NER genes needs to be determined in future studies.

A CSB/ERCC6 defect was first described in CSB, a rare UV-sensitive human disorder associated with a defective transcription-coupled repair pathway (11). CSB mice are more susceptible to chemically induced skin cancer than are wild-type and heterozygous mice (12) due to mutation of the CSB/ERCC6 gene, which influences and inactivates the coupling of transcription and repair, and results in the loss of normal rapid repair (13) and slow removal of photolesions from the transcribed strands of active genes (14). Studies using cells derived from CS or XP (G group) patients have demonstrated that these cells are deficient in the preferential repair of bulky adducts and ionizing radiation-induced DNA damage (15). Additionally, XPG protein is a key component in NER and encodes a DNA endonuclease involved in making an incision 3' to DNA lesions during removal of a damaged oligonucleotide (16). In summary, a defect in these genes may increase the susceptibility to cancer by allowing unrepaired DNA damage to remain in place, thus leading to carcinogenesis. However, the findings need to be confirmed by larger studies.

There are ubiquitous environmental toxicants and endogenous metabolites that can cause DNA damage, and apparently normal individuals with low levels of expression of DNA repair genes may have a low DNA repair capacity. If cellular DNA repair activity increases in response to DNA damage, transcription and translation activities should also be expected to increase concomitantly. The results of the present study provide evidence that the wide variation in expression level of NER genes reflects the existence of genetically determined factors that may contribute to interindividual variation in susceptibility. However, we cannot rule out the possibility that the reduced NER gene expression pattern in lung cancer patients may also be a consequence of an altered or impaired regulatory role of tumor suppressor genes such as p53 and other related genes (17,18). For example, p53 has been shown to bind to single-stranded DNA and interact in vitro and in vivo with NER proteins ERCC3, XPB, XPD and CSB (19,20), suggesting that it could have an effect on NER and other repair activities. Another possible mechanism for transcriptionally regulated gene expression is DNA methylation status. For instance, the hypermethylation of the promoter regions of O6-methylguanine-DNA methyltransferase (MGMT) (21) and hMLH1 (22) result in reduced expression of these genes. This epigenetic mechanism may exist and influence the expression...
of NER genes without genetic alterations, although little is known about the effect of methylation on the expression of these NER genes tested in the present study.

It has been reported that DNA repair capacity varies from individual to individual (3,4,23). Specifically, our observation of reduced expression level of NER genes in lung cancer patients supports the hypothesis that reduced expression of NER genes may reflect global DNA repair capacity and consequently modulates the risk of lung cancer. It appears that NER gene expression measured by the multiplex RT–PCR assay can provide additional information on the possible involvement of specific DNA repair genes in reduced DNA repair capacity. In this regard, although our findings are preliminary, the reduced expression level of NER genes in lung cancer patients as demonstrated in the present study warrants large-scale, well-designed epidemiological studies to further assess the role of DNA repair gene expression in the etiology of lung cancer.

Although it may be a valuable tool for large-scale epidemiological studies, this multiplex RT–PCR does not provide absolute quantitation of transcripts other than a semi-quantitative measurement of the relative RNA levels. More precise methods are needed to reduce assay variation in the measurements. The high-throughput technology of real-time PCR for quantitative gene expression (24) and microarray for measuring expression of a large number of genes at a time (25) will revolutionize the methods for detection of gene expression (RNA) in large epidemiological studies.

Acknowledgements

We thank Dr Reuben Lotan for critical review of the manuscript, Dr Don Norwood (Department of Scientific Publications) for editing the manuscript, Ms Susan Honn for patient recruitment, Mr Wayne Gosbee for data management, and Ms Yongli Guan and Mr Zhaozheng Guo for technical assistance. This study was supported in part by National Institutes of Health grants RO3 CA78425 (L.C.), RO1 CA55769 (M.R.S.), and RO1 CA74851 and R29,70334 (Q.W.). W.K.H. is an American Cancer Society Research Professor.

References