Expression of Nucleotide Excision Repair Proteins in Lymphocytes as a Marker of Susceptibility to Squamous Cell Carcinomas of the Head and Neck

Qingyi Wei,1,4 Li-E Wang,1 Erich M. Sturgis,1,2 and Li Mao3,4

Departments of 1Epidemiology, 2Head and Neck Surgery, and 3Thoracic/Head and Neck Medical Oncology, The University of Texas, M.D. Anderson Cancer Center; and 4The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas

Abstract

The transcript levels of nucleotide excision repair (NER) genes were shown to be associated with risk of squamous cell carcinomas of the head and neck (SCCHN). However, this association may be biased, because the transcript level does not necessarily reflect the level of protein expression. To address this issue, we did a pilot study to test the hypothesis that the expression of six core NER proteins is associated with risk of SCCHN. We obtained cultured lymphocytes from 57 patients with newly diagnosed SCCHN patients and 63 cancer-free controls. We transfected some of the lymphocytes with both damaged and undamaged plasmid DNA and quantified NER protein levels in these lymphocytes using a reverse-phase protein microarray. The relative NER protein levels in the 63 controls were highly correlated with each other (P < 0.001 for all). Compared with the controls, the cases had lower expression levels for all the NER proteins, particularly XPC and XPF, which were reduced by about 25% (P < 0.01). When we used the median expression levels of the NER proteins in the controls as cutoff values, we found that a significantly increased risk of SCCHN was associated with low expression of XPA (odds ratio (OR), 2.99; 95% confidence interval (CI), 1.22-7.47), XPC (OR, 2.46; 95% CI, 1.04-5.87), XPD (OR, 3.02; 95% CI, 1.18-7.76), and XPF (OR, 5.29; 95% CI, 2.01-13.9), but not ERCCI and XPG, after adjustment for age, sex, ethnicity, smoking, alcohol use, and sample storage time. In a multivariate logistic regression model that included all covariates and NER proteins, however, only low expression of XPF remained a significant risk factor for SCCHN (OR, 11.5; 95% CI, 2.32-56.6). These results suggest that XPF may be a crucial rate-limiting factor in DNA repair and that the reverse-protein microarray assay may be a useful tool for measuring protein markers of susceptibility to cancer. (Cancer Epidemiol Biomarkers Prev 2005;14(8):1961–6)

Introduction

Squamous cell carcinomas of the head and neck (SCCHN) are common malignancies, with >500,000 new cases worldwide estimated each year (1). In the U.S. in 2004, there were ~37,200 new cases of and 11,000 deaths from SCCHN (2). Many factors contribute to SCCHN, including tobacco smoking (3), alcohol use (4), viral infection (5), and genetic factors (6). Although smoking and alcohol use have a major role in the etiology of SCCHN, only a fraction of smokers and drinkers develop SCCHN, suggesting interindividual variations in genetic susceptibility to SCCHN in the general population.

Cellular DNA is constantly damaged by various endogenous and exogenous agents, including the recognized DNA adduct–inducing carcinogens contained in tobacco smoke. Sophisticated DNA repair pathways and mechanisms have evolved to maintain genomic integrity after insults from environmental hazards. One of the most important of these DNA repair pathways is nucleotide excision repair (NER; ref. 7). A number of crucial proteins, including seven core factors (ERCCI, XPA, XBP, XPC, XPD, XPF, and XPG), participate in NER (7). Functional mutations in any one of the seven genes encoding these core factors can lead to abnormal NER and thereby increase susceptibility to cancer (8). Several rare syndromes are characterized by NER deficiency coupled with high sensitivity to UV light and increased risk of cancer (9). Patients with xeroderma pigmentosum, for example, have mutations in at least one of seven NER genes and are extremely sensitive to sunlight-induced skin damage. Consequently, these patients have very high incidences of nonmelanoma skin cancer, melanoma, and other solid tumors (9).

In previous studies, we showed that an increased risk of SCCHN is associated with reduced DNA repair capacity, as measured by the host-cell reactivation assay (10), and with reduced levels of NER mRNA in lymphocytes (11). However, the transcript levels may not accurately reflect the level expression of proteins that perform the repair functions. To test the hypothesis that reduced expression of NER proteins is associated with increased risk of SCCHN, we developed a proteomic microarray assay to measure NER protein expression in lymphocytes from SCCHN patients and cancer-free controls.

Materials and Methods

Sample Procurement. The research protocol for this study as a part of an ongoing large molecular epidemiology of SCCHN was approved by the Institutional Review Board of The University of Texas M.D. Anderson Cancer Center. We used previously cryopreserved, viable peripheral blood lymphocyte samples from 57 patients in an ongoing case-control study of SCCHN. The sample collection started in 2000 with patients who had newly diagnosed, untreated SCCHN, that was histologically confirmed at The University of Texas M.D. Anderson Cancer Center; and 4The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas.

Received 2/6/05; revised 5/20/05; accepted 6/6/05.

Grant support: NIH grants R01 ES11740, CA 97007, PO1 CA106451, CA100264, and CA16672, and Department of Defense grant DAMD 17-03-1-0706.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

Requests for reprints: Qingyi Wei, Department of Epidemiology, Unit 189, and Li Mao, Department of Thoracic & Head and Neck Medical Oncology, Unit 432, The University of Texas, M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-3020; Fax: 713-563-0999. E-mail: qwei@mdanderson.org.

Copyright © 2005 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-05-0101
of Texas M.D. Anderson Cancer Center. The 63 cancer-free controls were frequency-matched with cases on age (+5 years), sex, and ethnicity that were obtained from a structured questionnaire.

We selected those subjects whose cryopreserved samples contained sufficient lymphocytes for cell culture and subsequent transfection with plasmid DNA damaged by benzo(a)pyrene diol epoxide, a tobacco carcinogen we previously used for studying host-cell NER DNA repair capacity as measured by the host-cell reactivation assay (12). The blood sample processing, plasmid preparation, and transfection have been described in detail previously (12). Briefly, lymphocytes were isolated from whole peripheral blood by Ficoll gradient centrifugation, cryopreserved within 24 hours with freezing medium, and stored in a −80°C freezer in 1.5 mL aliquots.

Cell Culture and Protein Preparation. The cryopreserved cells in each vial were quickly thawed and mixed, before the last trace of ice disappeared, with 8.5 mL of thawing medium (50% fetal bovine serum, 40% RPMI 1640, and 10% dextrose; purchased from Sigma Chemical, St. Louis, MO). This thawing method ensured a cellular viability of >80%, as confirmed by exclusion with 0.4% trypan blue (Sigma). After being washed twice with the thawing medium, the cells were incubated in RPMI 1640 (Life Technologies, Grand Island, NY) supplemented with 20% fetal bovine serum (Life Technologies) and stimulated with 56.25 μg/mL phytohemagglutinin (Murex Diagnostics, Norcross, GA) at 37°C for 72 hours. Only stimulated lymphocytes were expected to uptake the plasmids (13) and have active NER (14, 15).

After stimulation, the cells (~1 x 10^6) were collected and transfected by the DEAE-dextran (Pharmacia Biotech, Piscataway, NJ) method with 0.25 μg of untreated plasmids (as the baseline for comparison) or benzo(a)pyrene diol epoxide–damaged plasmids. In keeping with the protocol for the host-cell reactivation assay, in which the reactivation of the report gene is measured by quantifying the enzyme activity (12). The cells were collected for protein extraction 40 hours after the transfections. This procedure is crucial to ensure that the repair process is activated by the presence of the damaged plasmids, which served as the substrate for the NER enzymes. Thirty microliters of cell suspension (~1 x 10^6 cells) from each patient sample was mixed with 10 μL of 4× SDS sample buffer containing 50 μg/mL Tris-Cl (pH 8.0), 150 mmol/L NaCl, 0.1% SDS, and 1% Triton X-100 supplemented with a protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN). The cell lysate was then boiled for 5 minutes and stored at −80°C.

Construction of Reverse-Protein Microarrays. Proteins were extracted from the cells and were used to construct the microarrays. The extracted protein samples were serially diluted 1:1 with PBS (pH 7.5) to achieve final total protein concentrations ranging from 1 to 0.025 μg/μL. The minimum detection concentration was 0.0525 μg/μL. The serial dilutions were applied to FAST slides (Schleicher & Schuell Bioscience, Keene, NH) using a SpotBot microarrayer (TeleChem International, Cupertino, CA). Each sample containing the antigens (the NER proteins) to be detected was spotted in duplicate. Prepared slides were either used immediately or stored at −20°C.

Quantitative Analysis of Protein Levels Using Reverse-Protein Microarrays. We used mouse anti-human monoclonal or anti-goat or anti-rabbit polyclonal antibodies against XPD and XPG (Santa Cruz Biotechnology, Santa Cruz, CA); XPA, XPC, and XPF (Abcam, Cambridge, MA); ERCC1 (Novus Biological, Littleton, CO); and β-actin (Sigma). XPB was not assayed because no commercially available antibodies against XPB were specific enough for this study.

Briefly, the protein-spotted slides were treated with ReBlot (Chemix, Temecula, CA) for 15 minutes and then washed twice for 10 minutes each with washing buffer containing 300 mmol/L NaCl, 0.1% Tween 20, and 50 mmol/L Tris (pH 7.6). The protein arrays were then blocked with 1-Block (Applied Biosystems, Foster City, CA) for 30 minutes at room temperature. The primary antibodies were diluted based on their affinities, which were determined in our preliminary tests (data not shown). The dilution ratios were 1:300 for XPA and ERCC1; 1:500 for XPC, XPD, XPF, and XPG, and 1:100,000 for β-actin. The arrays were incubated with individual antibodies for 1 hour at room temperature. The anti-mouse, anti-goat, and anti-rabbit secondary antibodies (Vector Laboratories, Burlingame, CA) were labeled with biotin, diluted 1:30,000, and added to the slides, which were then incubated at room temperature for 30 minutes. Signals were enhanced using a catalyzed signal amplification system (DAKO, Carpinteria, CA) according to the manufacturer’s protocol, except that in the final step, we incubated the slides with Cy5-conjugated streptavidin (1:1,000; Jackson ImmunoResearch Laboratories, West Grove, PA) for 30 minutes. After each incubation step, the arrays were washed thrice for 5 minutes each with the washing buffer described above.

Signals on the protein microarrays were scanned on a ScanArray Lite microarray scanner (Perkin-Elmer Life Sciences, Boston, MA). The signal intensity of each spot and its background signal were analyzed using a ScanArray Express 2.0 microarray analysis system (Perkin-Elmer Life Sciences) running the ‘Run-easy Quant’ protocol. The final data were stored as graphic images for further analysis. Any scan-reading value <2,000 was treated as missing data. The median value of the scan-reading data for each dot of a protein on the microarray was used to calculate the means of the duplicates. Scan-reading data for the β-actin were used as a baseline to obtain the relative expression levels for each protein. The coefficient of variation (CV) was calculated as [SD / (mean) x 100].

Statistical Analysis. The mean signal intensity values for the cases and controls for each NER protein were compared with Student’s t test. The distribution of select variables of the cases and controls were compared with the χ² test. The correlation between the expression levels of different proteins was analyzed by the Pearson correlation coefficient. The median protein expression for the controls was used as the cutoff value for calculating the odds ratios (OR) associated with low expression and their 95% confidence intervals (CI). Multivariate logistic regression models were used to calculate the adjusted ORs and 95% CIs with adjustment for age (in years), sex (male versus female), ethnicity (non-Hispanic versus others), and sample storage time (in months). P values <0.05 were considered statistically significant. All statistical analyses were performed with SAS software version 8.0e (SAS Institute, Cary, NC).

Results

Protein Microarray Data. We began by testing the reproducibility and linearity of the reverse-protein microarray assay. Four samples of cell extracts from four controls were each diluted 1:1 five times; all 20 aliquots were spotted in triplicate on one slide (a total of 60 spots). We made three such slides for each sample and probed them with antibodies to XPA and XPF (Fig. 1A). The expression levels of XPA and XPF were linear on a log scale at the tested total protein concentrations between 1.0 and 0.0525 μg/μL (Table 1). These results were consistent between all samples and repeated experiments (Fig. 1B). Based on the CVs, however, it seemed that the reproducibility of the results was better at higher protein concentrations (CV as low as 0.8%) than at lower concentrations (CV as high as 24.6%; Table 1).

## Table 1
<table>
<thead>
<tr>
<th>Protein</th>
<th>Median CV</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPA</td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>XPF</td>
<td>0.6%</td>
<td></td>
</tr>
<tr>
<td>XPB</td>
<td>1.2%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Reproducibility and linearity of reverse-protein microarray assay data. A. Serial dilutions of four cell-extract samples (1, 2, 3, and 4) were spotted in triplicate and probed with specific antibodies against XPA and XPF. Each slide was tested three times (repeats 1, 2, and 3). The statistical data are summarized in Table 1. B. The mean intensities of the spots (in arbitrary units) were plotted on a log scale against the number of dilutions of the cell extract. C. Reverse-protein microarray assay of expression of the NER proteins ERCC1, XPA, XPC, XPD, XPF, and XPG. The level of β-actin was used as an internal control for standardization of the expression levels. The statistical data are summarized in Table 3.
We next diluted each of the 120 test samples twice, and then spotted them on the arrays in duplicate (Fig. 1C). The means of the duplicate readings were used to calculate the relative protein expression. Because of mechanical problems during the spotting, some protein spots were unreadable on the arrays, consequently, valid readings could not be obtained for up to five samples per protein. Consistent with the results of our reproducibility and linearity tests, the readings from spots with lower protein concentrations had greater variation, whereas the spots with higher protein concentrations produced consistent, strong, and readable signals. We therefore used the latter data (i.e., the original samples without dilution) to calculate the relative expression levels.

We also transfected cells from these samples with either undamaged plasmids or benzo(a)pyrene diol epoxide–damaged plasmids to stimulate DNA repair activity. When we compared the relative expression of NER proteins between the cases and controls, we found that the data from the samples transfected with damaged plasmids were a better predictor of risk of SCCHN (data not shown), although the two data sets were statistically correlated (P < 0.01). We therefore used data derived from cells transfected with damaged plasmids in the following experiments.

Subject Characteristics. Our analysis included 57 patients with newly diagnosed SCCHN and 63 cancer-free controls whose cryopreserved lymphocytes were available for culture, transfection, and protein extraction. The cases and controls were frequency-matched on age, sex, and ethnicity. The cases were slightly younger (56.2 versus 57.2 years) and comprised more males and non-Hispanic whites than did the controls, but these differences were not statistically significant (Table 2).

There were more smokers and alcohol drinkers among the cases than among the controls, and these differences were statistically significant (Table 2). Because the cases were recruited before the controls, the duration of lymphocyte storage was also significantly different between the two groups.

**Table 1. Reproducibility and linearity of reverse-protein microarray data**

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>XPA</th>
<th>XPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total protein (µg/µL)</td>
<td>0.96</td>
<td>0.48</td>
</tr>
<tr>
<td>Signal intensity (arbitrary units)</td>
<td>4,641</td>
<td>2,375</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>36,356</td>
<td>19,797</td>
</tr>
<tr>
<td>CV (%)</td>
<td>79</td>
<td>12</td>
</tr>
</tbody>
</table>

**Table 2. Characteristics SCCHN patients and control subjects**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cases (n = 57)</th>
<th>Controls (n = 63)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤55</td>
<td>24 (42)</td>
<td>25 (40)</td>
<td>0.787</td>
</tr>
<tr>
<td>&gt;55</td>
<td>33 (58)</td>
<td>38 (60)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>0.112</td>
</tr>
<tr>
<td>Male</td>
<td>43 (75)</td>
<td>39 (62)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>14 (25)</td>
<td>24 (38)</td>
<td></td>
</tr>
<tr>
<td>Ethnic group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>3 (5)</td>
<td>3 (5)</td>
<td>0.692</td>
</tr>
<tr>
<td>Mexican-American</td>
<td>3 (5)</td>
<td>4 (6)</td>
<td></td>
</tr>
<tr>
<td>Non–Hispanic White</td>
<td>51 (90)</td>
<td>56 (89)</td>
<td></td>
</tr>
<tr>
<td>Tobacco use†</td>
<td>42 (74)</td>
<td>31 (49)</td>
<td>0.006</td>
</tr>
<tr>
<td>Alcohol use†</td>
<td>43 (75)</td>
<td>33 (52)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>14 (25)</td>
<td>30 (48)</td>
<td></td>
</tr>
</tbody>
</table>

*Calculated using two-sided χ² test.
† The question asked was, “Have you ever smoked at least 100 cigarettes in your lifetime?”
‡ The question asked was, “Have you ever drunk alcoholic beverages at least once a week for 1 year or more?”
Obtained from logistic regression model with adjustment for age, sex, race, smoking status, alcohol use, and lymphocyte storage time.

*Expression level relative to that of β-actin.

Discussion

Our reverse-protein microarray assay successfully detected the target proteins at a total protein concentration as low as 0.025 µg/µL. However, the measurements seemed to be more reproducible at a total protein concentration ≥0.5 µg/µL. The cell extract from ~1 x 10⁶ cells (yielding 30 µL of sample) would thus be sufficient for repeated experiments, because each printed spot contained only 0.0033 µL. Using this assay, we showed that the relative expression levels of the six NER proteins (ERCC1, XPA, XPC, XPD, XPF, and XPG) were consistently significantly lower among the SCCHN patients than among the controls. Four of the six NER proteins tested (XPA, XPC, XPD, XPF) were associated with a significantly increased risk of SCCHN.

The data from this study are consistent with those in two of our previously published studies (10, 11). In the first study, we measured DNA repair capacity in 55 newly diagnosed SCCHN patients and 61 controls by the host-cell reactivation assay using a benzo(a)pyrene diol epoxide–damaged reporter gene (10). The mean DNA repair capacity in that study was significantly lower in the cases than in the controls. Those levels of these NER proteins were highly correlated with each other, the relative expression levels of all proteins were simultaneously adjusted for each other in the final multivariate logistic regression model containing age, sex, ethnicity, smoking, alcohol use, and sample storage time. The only significant adjusted OR was for XPF (11.5; 95% CI, 2.32-56.6) in the presence of other proteins in the same model (Table 4).
with DNA repair capacity values in the middle and lowest tertiles had >2-fold and 4-fold increased SCCHN risk, respectively, compared with those whose DNA repair capacity values were in the highest tertile. In the subsequent study, we investigated which NER genes might be responsible for the reduced DNA repair capacity in SCCHN. We previously measured the relative expression of the genes encoding five NER proteins (ERCC1, XPB, XPG, CSB, and XPC) by a multiplex RT-PCR method (11). The relative mRNA expression levels of ERCC1, XPB, XPG, and CSB were significantly lower in the cases than in the controls, and the risk of SCCHN associated with low expression of these genes was higher by 2- to 6-fold (11). In that study, we were not able to measure the expression of XPA, XPD, or XPF because the sequences of the genes were unknown at that time and the high level of sequence homology in the genome for the primers chosen made the assays unreliable.

In the present study, simultaneous adjustment for the expression levels of all proteins and other confounding factors revealed that the relative expression level of XPF was the only independent risk factor for SCCHN. Low compared with high expression of XPF was associated with an SCCHN risk >11-fold higher. Although the estimate was imprecise as evidenced by the wide 95% CI, this finding suggests that XPF may play a role in the repair of carcinogen-damaged DNA. Because ERCC1 needs XPF to form a functional complex (7), it is possible that XPF acts as a rate-limiting modulator. Based on our data, ERCC1 was expressed at higher levels than the other five proteins were, whereas XPF expression was <70% of ERCC1 expression. It is possible that our system was saturated with ERCC1 protein, so the amount of XPF became crucial for modulating the overall DNA repair capacity.

The present study is an extension of our previous studies assessing the best biomarkers of DNA repair capacity for predicting susceptibility to SCCHN. In the present study, we measured the relative expression levels of six of the seven core NER proteins because we did not find an appropriate antibody for XPB. Our data further support the notion that altered NER proteins because we did not find an appropriate antibody for XPB. Our data further support the notion that altered NER proteins because we did not find an appropriate antibody for XPB. However, the enzyme system required for DNA excision repair. Cell 1992;69:367 – 74.

There are several advantages to the reverse-protein microarray assay. First, compared with the host-cell reactivation assay (12), the microarray assay requires significantly (3-fold) fewer viable lymphocytes for protein extraction. Second, compared with the RT-PCR assay, the microarray assay is highly sensitive and reproducible, which is optimal for large-scale screening. Third, the microarray assay has the potential to test virtually any protein involved in NER or other molecular pathways underlying increased cancer risk. Finally, the microarray assay is rapid and cost-effective and produces a large quantity of data. With the availability of antibodies for specific protein posttranslational modifications, the microarray method may also become a powerful tool to assess functional changes in proteins.

Although the design of this pilot case-control study has inherent limitations of recall and selection biases, the reverse-protein microarray assay may be a powerful tool for future prospective studies if it is technically fine-tuned and the sampling issues resolved (16, 17). For instance, future studies must address the differences in protein concentrations between surrogate and target tissues, between fresh and stored serum samples, and before and after cancer diagnosis and treatment. An improved reverse-protein microarray assay should become a useful tool for future hypothesis-driven molecular epidemiologic studies of cancer.

Acknowledgments

We thank Dr. Margaret R. Spitz for critical review; Margaret Lung and Dr. Peggy Schuber for assistance in recruiting the subjects; Youhong Fan, Zhaozheng Guo, andYawei Qiao for laboratory assistance; Betty Jean Larson and Joanne Sider for manuscript preparation; and Pierrette Lo for scientific editing.

References

Expression of Nucleotide Excision Repair Proteins in Lymphocytes as a Marker of Susceptibility to Squamous Cell Carcinomas of the Head and Neck

Qingyi Wei, Li-E Wang, Erich M. Sturgis, et al.