Nucleotide Excision Repair Proteins and Risk of Head and Neck Squamous Cell Carcinomas in a Chinese Population

Pengyu Ren
Xi'an Jiaotong University
Xiaorong Niu
Xi'an Jiaotong University
Ruimin Zhao
Xi'an Jiaotong University
Zichen Chen
Xi'an Jiaotong University
Hao Dai
Xi'an Jiaotong University
Ling Zhang
Xi'an Jiaotong University
Bo Kou
Xi'an Jiaotong University
Chongwen Xu
Xi'an Jiaotong University
Xiaobao Yao
Xi'an Jiaotong University
Baiya Li
Xi'an Jiaotong University
Yuan Shao
Xi'an Jiaotong University
Shaoqiang Zhang
Xi'an Jiaotong University
Yiqian Liang
Xi'an Jiaotong University
Peng Han (penghanent@hotmail.com)
Xi'an Jiaotong University https://orcid.org/0000-0003-4115-4439

Research article

Keywords: Head and neck cancer, NER proteins, Cancer susceptibility, DNA repair, RPPA assay

Posted Date: August 26th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-57356/v1

Abstract

Background Nucleotide excision repair (NER) is pivotal in the development of smoking-related malignancies. We hypothesize that expression levels of NER proteins are associated with risk of the head and neck squamous cell carcinomas (HNSCCs) in a Chinese population.

Methods To test this hypothesis, we conducted a case-control study of 337 HNSCC patients and 285 cancer-free controls by measuring the expression levels of nine core NER proteins in cultured peripheral lymphocytes.

Results Compared with the controls, cases had statistically significantly lower expression levels of XPA (P=0.001). After dividing the subjects by controls' medians of expression levels, we found an association between an increased risk of HNSCCs and low XPA expression levels [adjusted ORs and 95% CIs:1.42 and 1.03-1.96; P_{trend}=0.031]. We identified a multiplicative interaction between smoking as well as drinking status and XPA expression levels (P = 0.005 and 0.044, respectively). Finally, the sensitivity of the expanded model with protein expression levels, in addition to demographic variables, on HNSCCs risk was significantly improved, especially among ever smokers and ever drinkers.

Conclusions Reduced XPA expression levels were associated with an increased risk of HNSCCs in a Chinese population.

Background

Head and neck squamous cell carcinomas (HNSCCs) are among the most common malignancies worldwide, which originate in the epithelial cells of the mucosal linings of the upper airway and food passages (the oral cavity, oropharynx, hypopharynx and larynx)[1-4]. In the United States, the estimated number of new HNSCC cases has been increasing from 59,260 in 2010 to 65,410 in 2019, according to the American Cancer Society[5-7]. In China, there were 74,500 new HNSCC cases in 2015, according to the Chinese Cancer Society[8]. It is well established that tobacco smoking and excessive alcohol use are major risk factors for smoking-related HNSCCs and prior high-risk human papillomaviruses (HPVs) infection for HPV-positive HNSCCs, especially for the oropharyngeal cancer in the Western world[9-11]. Although these risk factors may contribute to HNSCCs, only a small fraction of those who have the history of smoking, excessive alcohol use or HPV infection develop one of these cancers in their lifetime, suggesting that there may be heterogeneity in HNSCC susceptibility[12, 13].

Numerous carcinogenic chemicals in cigarette smoke can cause damages to cellular DNA[14]. For example, one of these chemicals is benzo(a)pyrene diol epoxide (BPDE) which is found in cigarette smoke as well as in the environment as a result of fuel combustion. The BPDE can induce irreversible damage to DNA by forming DNA adducts to block transcription of essential genes[15, 16]. Several DNA repair processes have evolved to repair these damages, among which nucleotide excision repair (NER) is a major and well-studied mechanism[16-19]. NER essentially uses nine core proteins (XPA, XPB, XPC, XPD, XPF, XPG, ERCC1, DDB1 and DDB2) to restore the damaged DNA to normal one in living cells[20-22].

Functional mutations in any of these proteins may lead to abnormal NER and subsequently increase susceptibility to cancers including cancers of skin, lung, and head and neck, et al[23-27].

In a study of 57 HNSCC patients and 63 cancer-free controls, it is reported that an increased risk of HNSCCs was associated with reduced expression levels of NER proteins in lymphocytes in non-Hispanic Whites population[28]. Later, the same group validate these results in a study with a much larger sample size and more NER proteins[29]. Until now, there is no study exploring the above associations in Chinese population, in which the composition of HNSCCs is quite different from that in non-Hispanic Whites population. Specifically, oropharyngeal cancers account for the most of HPV-positive HNSCCs in the United States, and the HPV-positive oropharyngeal cancer cases in previous non-Hispanic Whites study were about 91.4% of all the oropharyngeal cancer cases, while in Chinese population the most of oropharyngeal cancer cases are HPV-negative, meaning they are primarily caused by cigarette smoking[29-33]. In addition, the etiology of smoking-related HNSCCs is different from that of HPV-positive HNSCCs[34-36]. Subsequently, a study from a different race group is required to further validate the previously reported association studies. Therefore, we conducted a case-control study to test associations between expression levels of nine core NER proteins and risk of HNSCCs in a Chinese population.

Methods

Study subjects

We recruited 337 HNSCC patients and 285 cancer-free controls from the First Affiliated Hospital of Xi'an Jiaotong University during the period between 2013 and 2018. The cases were selected based on the following criteria: 40 years and older, newly diagnosed, histologically confirmed HNSCCs but with no other cancers. The controls were recruited among visitors accompanying patients to the First Affiliated Hospital of Xi'an Jiaotong University; they were biologically unrelated to the cases, frequency-matched with cases by age and sex, and have no history of prior malignancies. The subjects included in currently study were all Chinese Han. A written informed consent was obtained from cases and controls. Participants who smoked more than 100 cigarettes during their lifetime were defined as ever smokers, of which those who had quit smoking at least one year were defined as former smokers and remaining was considered current smokers; others were considered never smokers. Participants who drank alcoholic beverages at least weekly for one year were considered as ever drinkers, of which those who had quit drinking more than one year were considered as former and the remaining was defined current drinkers; others were defined never drinkers. Each subject donated a 15-ml blood sample. The HPV status of all subjects were tested by RT-PCR assay. In the previous study, the expression levels of NER proteins were not correlated with the HPV status in non-Hispanic White population [29]. Since the number of the HPV-positive HNSCC cases were very limited with only three cases identified as HPV-positive, we could not infer that NER proteins expression were not correlated with the HPV status in current Chinese population. Thus, the HPV-positive HNSCC subjects were excluded to avoid further heterogeneity in current study. The

study protocol was approved by the First Affiliated Hospital of Xi'an Jiaotong University Institutional Review Board.

Reverse-phase Protein Lysate Microarrays

Details regarding the RPPA (reverse-phase protein lysate microarrays) assay have been reported previously[29]. In detail, we isolated T-lymphocytes from whole peripheral blood by Ficoll gradient centrifugation. Cellular proteins were extracted from the cells and prepared for the RPPA analysis. Serial diluted lysates applied to nitrocellulose-coated slides (Schleicher & Schuell BioScience, Inc., USA) by Aushon Arrayer (Aushon BioSystems, USA). Each sample containing the antigens (the NER proteins) to be detected was spotted in duplicate with additional positive and negative controls prepared from mixed cell lysates or dilution buffer, respectively. Each slide was probed with a validated primary antibody plus a biotin-conjugated secondary antibody. Mouse anti-goat or anti-rabbit polyclonal or anti-human monoclonal antibodies were used against XPA, XPB, XPC, XPD and ERCC1 (Santa Cruz, USA); XPF (Abcam, USA); XPG (Protein tech, USA); DDB1 and DDB2 (Invitrogen, USA). The arrays were incubated with individual antibodies for 1 h at room temperature. The secondary antibodies were added to the slides and incubated at room temperature for 30 min.

Signals were amplified using a Dako system according to the protocol as previously described[29]. We then incubated the slides with a secondary conjugated streptavidin for 30 min and observed the signals by DAB colorimetric reaction. The signals on the microarrays were processed using the Array-Pro Analyzer software (Media Cybernetics, USA) to determine spot intensity, which were then analyzed by a logistic model by the R package. A fitted curve was plotted with the relative log2 concentration of each protein on the X-axis and the signal intensities on the Y-axis using the B-spline model as previously described[37]. Protein concentrations were determined from the fitted curve for each lysate by the curve-fitting and normalized by the median value for protein loading as described[38, 39]. The RPPA_CF is the correction factor in RPPA. Samples were considered as an outlier, if the correction factor was below 0.25 or above 2.5.

Statistical Analysis

The distribution of demographic variables was evaluated between cases and controls by the Chi-square test. The differences in the relative expression levels of NER proteins were compared by Wilcoxon rank-sum test between cases and controls.

The medians of expression values were used in the controls as the cutoff values for calculating crude odds ratio (OR) and their 95% confidence intervals (CI). The associations between protein expression levels and HNSCC risk were estimated by computing ORs and CIs from multivariate logistic regression models. Further stratification analyses were used to evaluate effect modification of related NER protein expression levels and demographic variables. A multiplicative interaction was defined as when $OR_{11} > OR_{01} \times OR_{10}$, in which OR_{11} was the OR when both factors were present, OR_{10} was the OR when only factor 1 was present, and OR_{01} was the OR when only factor 2 was present[40].

To assess the effects of protein expression levels on HNSCC risk prediction, two risk models were constructed to examine the area under the receiver operating characteristic (ROC) curve (AUC): the baseline model including only demographic variables, and the protein model including the expression levels in addition to these demographic variables. All tests were two-sided, and $P < 0.05$ was considered significant. All statistical analyses were performed using SAS software (version 9.4; SAS Institute, Inc., Cary, NC).

Results

Characteristics of the Study Population

The summary of the distributions of selected characteristics of cases and controls is presented in Table 1. There were no significant differences in the distributions of age and sex between cases and controls. The average age was 58.8 years for the cases (median, 58; range, 40-91) and 58.7 years for the controls (median, 59; range, 40-90). Of all the subjects, 66.2% of cases and 64.2% of controls were male, 56.7% of cases and 50.1% of controls were ever smokers and 67.3% of cases and 61.7% of controls were ever drinkers. There were more current smokers (32.1%) and current drinkers (33.8%) in cases than in controls (17.5% and 31.9%, respectively). The primary HNSCCs of 337 patients included the oral cavity (67, 19.9%), oropharynx (150, 44.5%), hypopharynx and larynx (120, 35.6%).

Differences in NER Protein Expression Levels between Cases and Controls

The cases showed lower relative mean expression levels in six of the nine core NER proteins analyzed than did controls, except for XPC, XPG, and ERCC1 (Table 2). In Wilcoxon rank-sum test for differences in NER protein expression levels between cases and controls, only XPA levels were statistically significantly lower in cases than in controls (P= 0.001; Fig 1A). Because the expression levels of the nine NER proteins were measured at the same time, they were likely to be correlated with each other. As shown in Supplementary Table 1, expression levels of XPA were statistically significantly correlated with XPB, XPC, XPD, and ERCC1(P = 0.019, P = 0.050, and $P < 0.001$, and P = 0.012, respectively).

Stratification Analyses of Expression Levels of XPA by Selected Variables

Stratification analyses of XPA expression levels revealed that patients in subgroups of the age $\leq$ 59, age > 59, male, female, former and current smokers, and former and current drinkers exhibited significantly lower mean expression levels of XPA than did controls (All the $P < 0.001$, respectively, Table 3). In cases, women had lower expression levels of XPA than did men, but in controls, women had higher expression levels of XPA than did men, and the sex differences in the expression levels were insignificant in both case and control groups (P = 0.249 and P = 0.889, respectively, Table 3). Moreover, both ever smokers and drinkers had significant lower expression levels of XPA than did never smokers and drinkers, respectively (All the $P < 0.001$, respectively, Table 3). There were no significant differences in the expression levels of XPA by tumor sites, suggesting that expression levels of XPA may not be different among tumors of HNSCCs (Supplementary Table2).

Associations between NER Protein Expression Levels and Risk of HNSCCs

To estimate HNSCC risk, the relative expression levels were grouped into median values of the controls (Table 4). The crude ORs for HNSCC risk associated with lower relative expression levels of XPA were 1.43 (95% CI, 1.04-1.97), compared with the high expression levels of XPA. After adjusting for age, sex, smoking status and alcohol consumption in multivariate logistic regression analysis, the OR of XPA remained essentially unchanged. When continuous expression values were used in the logistic regression model with adjustment for all covariates, there was also a dose-response relationship between the reduced expression levels of XPA and the increased HNSCC risk (P_{trend}= 0.031).

Interactions between XPA Expression Levels and Selected Variables

We further assessed possible interactions on a multiplicative scale between expression levels of XPA and selected variables listed in Table 1. The multiplicative interaction was tested when we included the interaction term (i.e., relative expression levels of XPA × each of the risk factors) in a multivariate regression model that also included the main effects of NER protein expression levels and other covariates. We found that smoking status as well as drinking status had significantly multiplicative interactions with relative expression levels of XPA (P = 0.005 and P = 0.044, respectively, Table 3), in association with HNSCC risk. To further unravel these multiplicative interactions, we stratified the adjusted ORs by smoking status and drinking status. It was apparent that ORs for the relative expression levels of XPA by median in groups of ever smokers were greater than those of never smokers (Fig. 1B). And the ORs for the relative expression levels of XPA by medians in groups of ever drinkers were greater than those of never drinkers (Fig. 1C).

We further assessed the prediction performance of models integrating demographic variables and protein expression levels on HNSCCs using the ROC curves that measure the effect of XPA expression levels in two dimensions. The AUC was significantly improved in the model that included the effect of XPA expression levels, compared with the model that did not (Fig. 2A**,** P = 0.004). Furthermore, the AUC was significantly improved in former and current smokers that included the effects of XPA expression levels, compared with the model that did not (Fig. 2C and 2D, P < 0.001 and P < 0.001, respectively), but insignificantly improved in never smokers (Fig. 2B, P = 0.462). The AUC was significantly improved in former and current drinkers that included the effects of XPA expression levels, compared with the model that did not (Supplementary Fig. 1B and 1C, P = 0.001 and P = 0.001, respectively), but insignificantly improved in never drinkers (Supplementary Fig. 1A, P = 0.404).

Discussion

In this study, we further confirmed the previous study's results that reduced NER protein expression was associated with an increased risk of HNSCCs by the RPPA assay. Our results showed that the reduced relative expression levels of XPA were associated with an increased risk of HNSCCs in a Chinese population. We further assessed interactions between XPA expression levels and selected variables and

found that smoking as well as drinking had significant multiplicative interactions with XPA expression on HNSCC risk. Moreover, the AUC model suggested that the effects of XPA expression levels further improved the risk prediction in ever smokers and drinkers.

In an early study, it was reported that there was an association between an increased risk of HNSCCs and reduced expression levels of XPD, XPF, XPA and XPC in non-Hispanic population, when appropriate antibodies for DDB1 and XPB were not available at that time[28]. Later, the same group validated the above results with more available antibodies for essential proteins, and found the relative expression levels of XPA and XPB were significantly lower in cases than in controls, and the risk of HNSCCs associated with lower expression levels of XPA and XPB[29]. As the composition of HNSCCs in Chinese population is quite different from that in non-Hispanic Whites population, we tested the associations between expression levels of nine core NER proteins and risk of HNSCCs in a Chinese population, and found that the reduced expression levels of XPA was associated with HNSCC risk, but not for XPB. These results further support the notion that altered translational levels of NER genes, which have a more direct effect on the NER capacity than that of transcript levels, may contribute to the risk of HNSCCs. Moreover, our previous work of transcript level suggested that mRNA expression level of *XPA* and *XPB* were statistically significantly lower in cases than in controls, and the reduced mRNA expression levels of *XPB* were associated with an increased risk of HNSCCs in a Chinese population[41], however, we did not find the above association with *XPB* in translational level. One reason for this discrepancy is that the transcript levels and translational levels of NER genes may not be directly correlated. Although the mRNA of NER gene is ultimately translated into a NER protein, the transcription and translation processes are far from a simple linear correlation[42]. The underling mechanisms are likely to be the cis-acting and trans-acting processes create a serial of systems that promote or inhibit the synthesis of proteins from a certain copy number of mRNA molecules, and translation levels are more directly involved in the NER repair process[43]. Another reason is that the sample size of current study is still not large enough, future studies with more cases and controls are warranted to validate the current results.

Previous study suggested that a modification effect of smoking status on XPB, indicating that an association between the reduced expression levels of XPB and increased risk of HNSCCs may differ by smoking status[29]. In current study, we have observed smoking as well as drinking status had significant multiplicative interactions with XPA expression levels on HNSCC risk, other than XPB. Subsequently, we stratified the ORs of XPA by smoking and drinking status and found that the adjusted ORs for XPA in ever smokers or ever drinkers were greater than that in never smokers or never drinkers, indicating that ever smokers or ever drinkers might have a higher risk of developing HNSCCs with reduced XPA expression levels.

The XPA protein consists of several domains: the C-terminal domain able to interact with the transcription factor IIH, the N-terminal domain with RPA34 and ERCC1 binding sites, and the central domain responsible for DNA binding [44]. Variation in XPA's functions may lead to an aberrant NER process and subsequently increase the susceptibility to cancer. Our data suggested an increased risk of HNSCCs associated with reduced expression levels of XPA in a Chinese population, and the current results were

consistent with previously published non-Hispanic Whites studies on HNSCC risks, suggesting XPA may serve as a general biomarker for HNSCCs among two race groups.

Previously, we assessed the performance of NER proteins on HNSCC risk in the AUC model in non-Hispanic Whites population and found that the AUC model was significantly improved by including the combined effect of XPB and XPA expression, compared with the model that did not, especially in former smokers[29]. In current Chinese population study, we found that the AUC model was significantly improved by XPA expression levels, compared with the model that did not, especially in ever smokers and ever drinkers, suggesting that suboptimal XPA expression levels may play a more important role in the risk of HNSCCs in ever smokers and ever drinkers.

The RPPA assay is a rapid, cost-effective and most importantly an efficient method to measure the expression levels of NER proteins, and the current study is the first study to measure the associations between NER proteins and risk of HNSCCs in Chinese population. Although the present study is an extension of previous work for NER proteins with more antibody, there are still several limitations needed to be resolved. Although the results in current translational study is different from that of the transcriptional study, the current results in translation levels are more directly involved in the NER repair process. Like previous hospital-based studies, the control group may not be representative of the general population, and future studies may need a much larger sample size and recruit the controls from the community-based population.

Conclusion

Reduced XPA expression levels were associated with an increased risk of HNSCCs in a Chinese population. Future mechanistic studies are also needed for the role of XPA in the etiology of HNSCCs in a Chinese population.

Abbreviations

HNSCCs: head and neck squamous cell carcinomas;

NER: nucleotide excision repair;

RPPA: reverse-phase protein lysate microarrays;

HPVs: human papillomaviruses;

OR: odds ratio;

CI: confidence interval;

ROC: receiver operating characteristic curve

Declarations

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

We thank Hao Dai and Qiong Shi for her assistance in recruiting the subjects and gathering the questionnaire information.

Funding

This study was funded by Natural Science Foundation of China (No. 81803317, 81971766, 81902322), Natural Science Basic Research Plan in Shaanxi Province (No.2019JQ-957), and Fundamental Research Funds for the Central Universities, China (xjj2018094; xzy012019109).

Author information

Affiliations

Department of Otorhinolaryngology-Head and Neck Surgery, First Affiliated Hospital, Xi'an Jiaotong University College of Medicine, Xi'an, Shaanxi, 710061, PR China

Pengyu Ren, Xiaorong Niu, Ruimin Zhao, Ling Zhang, Bo Kou, Chongwen Xu, Xiaobao Yao, Baiya Li, Yuan Shao, Shaoqiang Zhang & Peng Han

Department of Neurosurgery, Second Affiliated Hospital, Xi'an Jiaotong University College of Medicine, Xi'an, Shaanxi, 710004, PR China

Pengyu Ren

Department of Otorhinolaryngology-Head and Neck Surgery, Second Affiliated Hospital, Xi'an Jiaotong University College of Medicine, Xi'an, Shaanxi, 710004, PR China

Zichen Chen

Department of Nuclear Medicine, First Affiliated Hospital, Xi'an Jiaotong University College of Medicine, Xi'an, Shaanxi, 710061, PR China

Yiqian Liang

Contributions

PH, YL- Conceptualization, Methodology.

ZC, LZ, Bo Ko-Performed the experiments

PR, BL, YS, HD- Analyzed the data

RZ, CX, XY-Code program

PR, SZ- Wrote the paper

Corresponding authors

Correspondence to Peng Han or Yiqian Liang.

Ethics declarations

Ethics approval and consent to participate

The study protocol was approved by the First Affiliated Hospital of Xi'an Jiaotong University Institutional Review Board. This study was performed in accordance with the Declaration of Helsinki. The consent obtained from study participants was written.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no conflicts of interest.

References

1. Wissinger E, Griebsch I, Lungershausen J, Byrnes M, Travers K, Pashos CL: **The humanistic burden of head and neck cancer: a systematic literature review**. *PharmacoEconomics* 2014, **32**(12):1213-1229.
2. Jou A, Hess J: **Epidemiology and Molecular Biology of Head and Neck Cancer**. *Oncology research and treatment* 2017, **40**(6):328-332.
3. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: **Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012**. *International journal of cancer* 2015, **136**(5):E359-386.
4. Brandstorp-Boesen J, Sorum Falk R, Boysen M, Brondbo K: **Impact of stage, management and recurrence on survival rates in laryngeal cancer**. *PloS one* 2017, **12**(7):e0179371.
5. Jemal A, Siegel R, Xu J, Ward E: **Cancer statistics, 2010**. *CA: a cancer journal for clinicians* 2010, **60**(5):277-300.
6. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2019**. *CA: a cancer journal for clinicians* 2019, **69**(1):7-34.

7. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2016**. *CA: a cancer journal for clinicians* 2016, **66**(1):7-30.
8. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J: **Cancer statistics in China, 2015**. *CA: a cancer journal for clinicians* 2016, **66**(2):115-132.
9. Kawakita D, Matsuo K: **Alcohol and head and neck cancer**. *Cancer metastasis reviews* 2017, **36**(3):425-434.
10. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C *et al*: **Human papillomavirus and survival of patients with oropharyngeal cancer**. *The New England journal of medicine* 2010, **363**(1):24-35.
11. Leemans CR, Snijders PJF, Brakenhoff RH: **The molecular landscape of head and neck cancer**. *Nature reviews Cancer* 2018, **18**(5):269-282.
12. Cancer Genome Atlas N: **Comprehensive genomic characterization of head and neck squamous cell carcinomas**. *Nature* 2015, **517**(7536):576-582.
13. Pytynia KB, Dahlstrom KR, Sturgis EM: **Epidemiology of HPV-associated oropharyngeal cancer**. *Oral oncology* 2014, **50**(5):380-386.
14. Xiong P, Hu Z, Li C, Wang LE, El-Naggar AK, Sturgis EM, Wei Q: **In vitro benzo[a]pyrene diol epoxide-induced DNA damage and chromosomal aberrations in primary lymphocytes, smoking, and risk of squamous cell carcinoma of the head and neck**. *International journal of cancer* 2007, **121**(12):2735-2740.
15. Li D, Wang LE, Chang P, El-Naggar AK, Sturgis EM, Wei Q: **In vitro benzo[a]pyrene diol epoxide-induced DNA adducts and risk of squamous cell carcinoma of head and neck**. *Cancer research* 2007, **67**(12):5628-5634.
16. Turgeon MO, Perry NJS, Poulogiannis G: **DNA Damage, Repair, and Cancer Metabolism**. *Frontiers in oncology* 2018, **8**:15.
17. Weaver AN, Cooper TS, Rodriguez M, Trummell HQ, Bonner JA, Rosenthal EL, Yang ES: **DNA double strand break repair defect and sensitivity to poly ADP-ribose polymerase (PARP) inhibition in human papillomavirus 16-positive head and neck squamous cell carcinoma**. *Oncotarget* 2015, **6**(29):26995-27007.
18. Sancar A: **Mechanisms of DNA Repair by Photolyase and Excision Nuclease (Nobel Lecture)**. *Angewandte Chemie* 2016, **55**(30):8502-8527.
19. Liu Z, Liu H, Gao F, Dahlstrom KR, Sturgis EM, Wei Q: **Reduced DNA double-strand break repair capacity and risk of squamous cell carcinoma of the head and neck--A case-control study**. *DNA repair* 2016, **40**:18-26.
20. Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K, Linn S: **Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints**. *Annual review of biochemistry* 2004, **73**:39-85.
21. Spivak G: **Nucleotide excision repair in humans**. *DNA repair* 2015, **36**:13-18.

22. Yuan H, Li H, Ma H, Niu Y, Wu Y, Zhang S, Hu Z, Shen H, Chen N: **Genetic polymorphisms in key DNA repair genes and risk of head and neck cancer in a Chinese population**. *Experimental and therapeutic medicine* 2012, **3**(4):719-724.
23. Cleaver JE: **Cancer in xeroderma pigmentosum and related disorders of DNA repair**. *Nature reviews Cancer* 2005, **5**(7):564-573.
24. Cleaver JE, Lam ET, Revet I: **Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity**. *Nature reviews Genetics* 2009, **10**(11):756-768.
25. Li C, Wang LE, Wei Q: **DNA repair phenotype and cancer susceptibility–a mini review**. *International journal of cancer* 2009, **124**(5):999-1007.
26. Park JY, Park SH, Choi JE, Lee SY, Jeon HS, Cha SI, Kim CH, Park JH, Kam S, Park RW *et al*: **Polymorphisms of the DNA repair gene xeroderma pigmentosum group A and risk of primary lung cancer**. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2002, **11**(10 Pt 1):993-997.
27. Hijazi H, Salih MA, Hamad MH, Hassan HH, Salih SB, Mohamed KA, Mukhtar MM, Karrar ZA, Ansari S, Ibrahim N *et al*: **Pellagra-like condition is xeroderma pigmentosum/Cockayne syndrome complex and niacin confers clinical benefit**. *Clinical genetics* 2015, **87**(1):56-61.
28. Wei Q, Wang LE, Sturgis EM, Mao L: **Expression of nucleotide excision repair proteins in lymphocytes as a marker of susceptibility to squamous cell carcinomas of the head and neck**. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2005, **14**(8):1961-1966.
29. Han P, Liu H, Shi Q, Liu Z, Troy JD, Lee WT, Zevallos JP, Li G, Sturgis EM, Wei Q: **Associations between expression levels of nucleotide excision repair proteins in lymphoblastoid cells and risk of squamous cell carcinoma of the head and neck**. *Molecular carcinogenesis* 2018, **57**(6):784-793.
30. Huang H, Zhang B, Chen W, Zhou SM, Zhang YX, Gao L, Xu ZG, Qiao YL, Tang PZ: **[Human papillomavirus infection in oropharyngeal squamous cell carcinoma and prognosis: a preliminary analysis of 66 cases]**. *Zhonghua er bi yan hou tou jing wai ke za zhi = Chinese journal of otorhinolaryngology head and neck surgery* 2012, **47**(3):207-211.
31. Zhang YX, Zhang B, Gao L, Xu ZG, Tang PZ: **[Clinical analysis of 318 cases of oropharyngeal squamous cell carcinoma]**. *Zhonghua er bi yan hou tou jing wai ke za zhi = Chinese journal of otorhinolaryngology head and neck surgery* 2013, **48**(5):398-404.
32. Shah A, Malik A, Garg A, Mair M, Nair S, Chaturvedi P: **Oral sex and human papilloma virus-related head and neck squamous cell cancer: a review of the literature**. *Postgraduate medical journal* 2017, **93**(1105):704-709.
33. Ducatman BS: **The Role of Human Papillomavirus in Oropharyngeal Squamous Cell Carcinoma**. *Archives of pathology & laboratory medicine* 2018, **142**(6):715-718.
34. Castellsague X, Alemany L, Quer M, Halec G, Quiros B, Tous S, Clavero O, Alos L, Biegner T, Szafarowski T *et al*: **HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of**

Biomarkers in 3680 Patients. *Journal of the National Cancer Institute* 2016, **108**(6):djv403.

35. Udager AM, McHugh JB: **Human Papillomavirus-Associated Neoplasms of the Head and Neck**. *Surgical pathology clinics* 2017, **10**(1):35-55.
36. Siddiqi K, Shah S, Abbas SM, Vidyasagaran A, Jawad M, Dogar O, Sheikh A: **Global burden of disease due to smokeless tobacco consumption in adults: analysis of data from 113 countries**. *BMC medicine* 2015, **13**:194.
37. Hu J, He X, Baggerly KA, Coombes KR, Hennessy BT, Mills GB: **Non-parametric quantification of protein lysate arrays**. *Bioinformatics* 2007, **23**(15):1986-1994.
38. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills GB, Kornblau SM: **Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells**. *Molecular cancer therapeutics* 2006, **5**(10):2512-2521.
39. Gonzalez-Angulo AM, Hennessy BT, Meric-Bernstam F, Sahin A, Liu W, Ju Z, Carey MS, Myhre S, Speers C, Deng L *et al*: **Functional proteomics can define prognosis and predict pathologic complete response in patients with breast cancer**. *Clinical proteomics* 2011, **8**(1):11.
40. Wang LE, Hu Z, Sturgis EM, Spitz MR, Strom SS, Amos CI, Guo Z, Qiao Y, Gillenwater AM, Myers JN *et al*: **Reduced DNA repair capacity for removing tobacco carcinogen-induced DNA adducts contributes to risk of head and neck cancer but not tumor characteristics**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2010, **16**(2):764-774.
41. Ren P, Niu X, Liu C, Liu J, Li H, Zhao Q, Xing J, Bai Y, Liang Y, Han P: **Associations between expression levels of nine core nucleotide excision repair genes in lymphocytes and risk of head and neck squamous cell carcinomas in a Chinese population**. *International journal of clinical oncology* 2020, **25**(4):660-669.
42. Greenbaum D, Colangelo C, Williams K, Gerstein M: **Comparing protein abundance and mRNA expression levels on a genomic scale**. *Genome biology* 2003, **4**(9):117.
43. Maier T, Guell M, Serrano L: **Correlation of mRNA and protein in complex biological samples**. *FEBS letters* 2009, **583**(24):3966-3973.
44. Wang M, Chu H, Zhang Z, Wei Q: **Molecular epidemiology of DNA repair gene polymorphisms and head and neck cancer**. *Journal of biomedical research* 2013, **27**(3):179-192.

Tables

Table 1. Distributions of demographic variables and tumor characteristics between cases and controls

Variable	Case (n = 337)	Control (n = 285)	*P**
Sex			0.609
Male	223 (66.2)	183 (64.2)	
Female	114 (33.8)	102 (35.8)	
Age			0.484
Median (range)	58 (40-91)	59 (40-90)	
≤ 59	188 (55.8)	151 (53.0)	
> 59	149 (44.2)	134 (47.0)	
Smoking			0.001
Never	146 (43.3)	142 (49.8)	
Former	83 (24.6)	93 (32.6)	
Current	108 (32.1)	50 (17.5)	
Drinking			0.331
Never	110 (32.6)	109 (38.3)	
Former	113 (33.5)	85 (29.8)	
Current	114 (33.8)	91 (31.9)	
Tumor site			
Oropharynx	150 (44.5)		
Larynx/Hypopharynx	120 (35.6)		
Oral cavity	67 (19.9)		

HNSCCs = head and neck squamous cell carcinomas

*Chi-square tests for the distribution comparison of the demographic variables between cases and controls

Table 2. Comparison of the expression levels of nine NER proteins between cases and controls

Protein	Mean ± SD		Median/IQR		Difference	*P*
	Case (n=337)	Control (n=285)	Case (n=337)	Control (n=285)	Case - Control	
XPA	0.199 ± 0.033	0.225 ± 0.066	0.201/0.044	0.215/0.023	-0.026	**0.001**
XPB	0.590 ± 0.066	0.591 ± 0.062	0.574/0.095	0.576/0.021	-0.001	0.152
XPC	0.222 ± 0.075	0.222 ± 0.078	0.220/0.101	0.220/0.095	0.000	0.606
XPD	0.393 ± 0.061	0.395 ± 0.083	0.361/0.070	0.363/0.071	-0.002	0.468
XPF	0.232 ± 0.047	0.239 ± 0.051	0.216/0.038	0.216/0.114	-0.007	0.175
XPG	0.386 ± 0.045	0.381 ± 0.050	0.389/0.061	0.381/0.057	0.005	0.051
ERCC1	0.250 ± 0.077	0.239 ± 0.074	0.231/0.108	0.231/0.102	0.011	0.078
DDB1	0.142 ± 0.020	0.143 ± 0.028	0.143/0.011	0.143/0.011	-0.001	0.630
DDB2	0.532 ± 0.192	0.540 ± 0.192	0.518/0.219	0.522/0.228	-0.008	0.663

SD = standard deviation; IQR = interquartile range

P value in Wilcoxon rank-sum tests; The relative concentrations of each protein were normalized for protein loading and transformed to linear values

Table 3. Stratification analyses of expression levels of XPA between cases and controls

Variable	XPA (Mean ± SD)		*P**	*P^*
	Case (n=337)	Control (n=285)		
Sex				0.951
Male	0.200 ± 0.032	0.223 ± 0.060	< 0.001	
Female	0.197 ± 0.035	0.227 ± 0.075	< 0.001	
*P**	0.249	0.889		
Age				0.196
≤ 59	0.200 ± 0.035	0.224 ± 0.065	< 0.001	
> 59	0.197 ± 0.030	0.226 ± 0.067	< 0.001	
*P**	0.796	0.999		
Smoking				0.005
Never	0.211 ± 0.028	0.226 ± 0.065	0.683	
Former	0.192 ± 0.029	0.218 ± 0.046	< 0.001	
Current	0.187 ± 0.035	0.233 ± 0.093	< 0.001	
*P***	< 0.001	0.301		
Drinking				0.044
Never	0.206 ± 0.029	0.224 ± 0.074	0.490	
Former	0.198 ± 0.035	0.224 ± 0.040	< 0.001	
Current	0.193 ± 0.033	0.227 ± 0.074	< 0.001	
*P****	0.001	0.246		

P value in Wilcoxon rank-sum tests

^*P* value in multiplicative interaction analysis between selected variables and proteins in relation to HNSCC risk

The relative concentrations of each protein were normalized for protein loading and transformed to linear values

Table 4. Logistic regression analysis of expression levels of nine NER proteins in cases and controls

NER Proteins	Protein Levels***	Case n (%)	Control n (%)	Crude OR (95% CI)	Adjusted OR* (95% CI)
XPA	≥ 0.215	139 (41.3)	143 (50.2)	1.00 (Ref)	1.00 (Ref)
	< 0.215	198 (58.7)	142 (49.8)	**1.43 (1.04-1.97)**	**1.42 (1.03-1.96)**
*P***				**0.026**	**0.031**
XPB	≥ 0.576	145 (37.1)	143 (50.2)	1.00 (Ref)	1.00 (Ref)
	< 0.576	192 (62.9)	142 (49.8)	1.33 (0.97-1.83)	1.33 (0.97-1.83)
*P***				0.075	0.075
XPC	≥ 0.220	173 (51.3)	142 (49.8)	1.00 (Ref)	1.00 (Ref)
	< 0.220	164 (48.7)	143 (50.2)	0.94 (0.69-1.29)	0.94 (0.69-1.29)
*P***				0.707	0.710
XPD	≥ 0.363	152 (45.1)	143 (50.2)	1.00 (Ref)	1.00 (Ref)
	< 0.363	185 (54.9)	142 (49.8)	1.23 (0.89-1.68)	1.23 (0.90-1.69)
*P***				0.207	0.199
XPF	≥ 0.216	168 (49.8)	143 (50.2)	1.00 (Ref)	1.00 (Ref)
	< 0.216	169 (50.2)	142 (49.8)	1.01 (0.74-1.39)	1.01 (0.73-1.38)
*P***				0.936	0.971
XPG	≥ 0.381	193 (57.3)	143 (50.2)	1.00 (Ref)	1.00 (Ref)
	< 0.381	144 (42.7)	142 (49.8)	0.75 (0.55-1.03)	0.75 (0.54-1.03)
*P***				0.077	0.074
ERCC1	≥ 0.231	195 (57.9)	142 (49.8)	1.00 (Ref)	1.00 (Ref)
	< 0.231	142 (42.1)	143 (50.2)	0.72 (0.53-1.06)	0.75 (0.52-1.06)
*P***				0.068	0.059
DDB1	≥ 0.143	173 (51.3)	144 (50.5)	1.00 (Ref)	1.00 (Ref)
	< 0.143	164 (48.7)	141 (49.5)	0.97 (0.71-1.33)	0.97 (0.71-1.33)
*P***				0.841	0.853
DDB2	≥ 0.522	167 (49.5)	142 (49.8)	1.00 (Ref)	1.00 (Ref)
	< 0.522	170 (50.5)	143 (50.2)	1.01 (0.74-1.39)	1.01 (0.73-1.38)

*P***	0.947	0.959

NER = nucleotide excision repair;

OR = odds ratio; CI = confidence interval

*Adjusted for age, sex, smoking and alcohol status

***P* value in trend test by continuous protein expression levels

***Expression levels by medians based on the median values of control subjects.

Figures

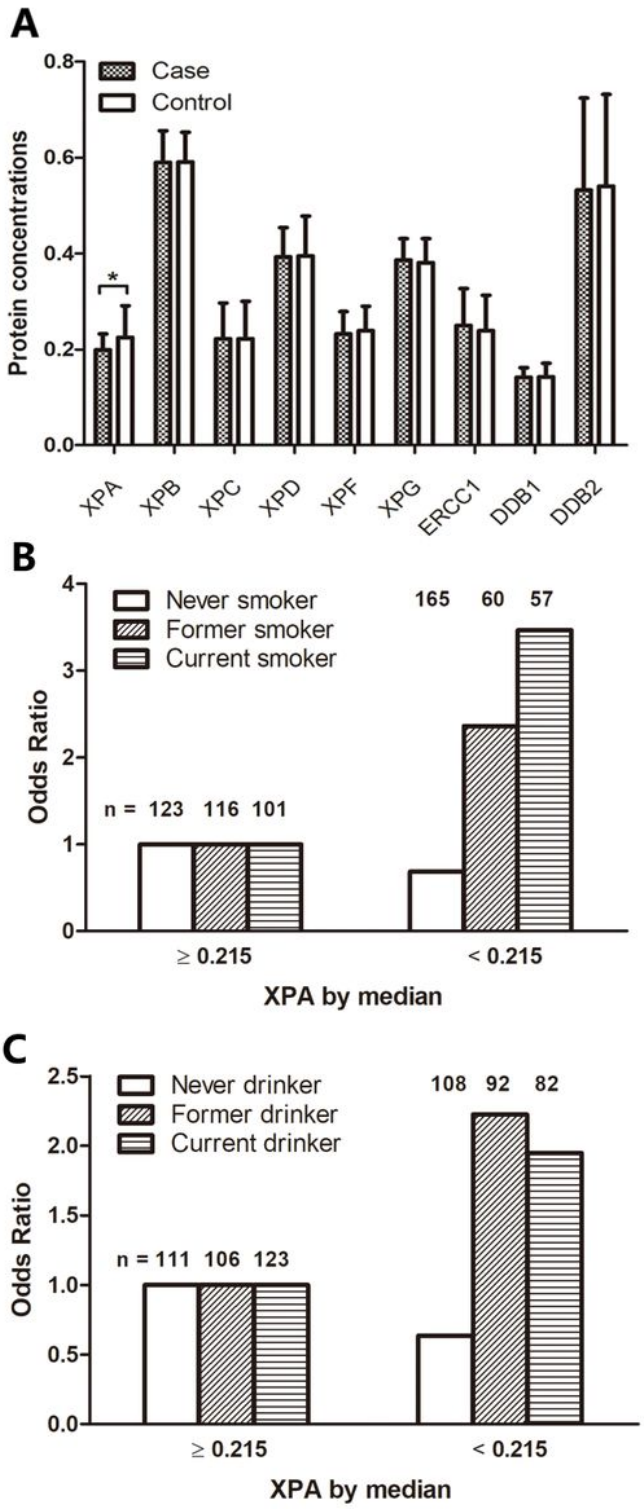

Figure 1

(A) Relative expression levels of nine NER proteins between HNSCC patients and healthy controls. Reverse-phase protein lysate microarrays were used to measure the relative expressions of nine NER proteins; (B) Modification effects of XPA by smoking status; (C) Modification effects of XPA by drinking status

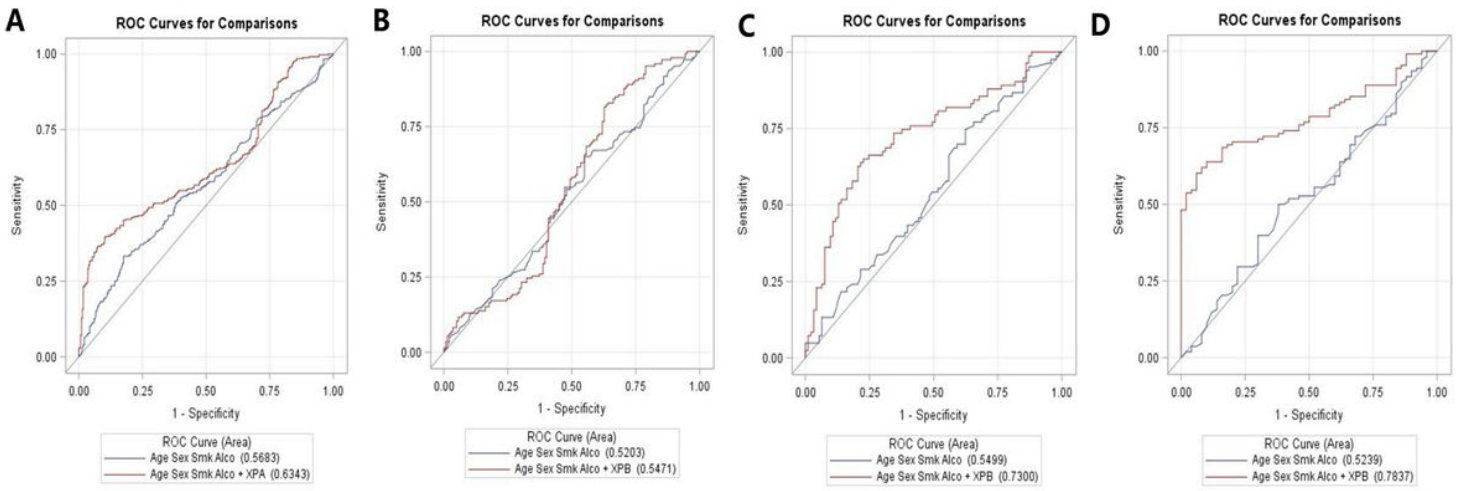

Figure 2

Overall and stratified ROC curves by smoking status calculated in multivariate logistic models. (A) The AUC was significantly improved in the model that included the effect of XPA expression levels, compared with the model that did not (P = 0.004); (B) The AUC was insignificantly improved in never smokers that included the effect of XPA expression levels (P = 0.462); (C) The AUC was significantly improved in former smokers (P < 0.001); (D) The AUC was significantly improved in current smokers (P < 0.001)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFig.docx
- SupplemantaryFIG.tif
- SupplementaryTables.docx