Published in final edited form as:
Clin Cancer Res. 2008 July 1; 14(13): 4345–4352. doi:10.1158/1078-0432.CCR-07-5282.

Functional Variants of the *NEIL1* and *NEIL2* Genes and Risk and Progression of Squamous Cell Carcinoma of the Oral Cavity and Oropharynx

Xiadong Zhai[1], **Hui Zhao**[1], **Zhensheng Liu**[1], **Li-E Wang**[1], **Adel K. El-Naggar**[2], **Erich M. Sturgis**[1,3], and **Qingyi Wei**[1]

[1] Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

[2] Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

[3] Department of Head and Neck Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Abstract

Purpose—Human DNA glycosylases NEIL1 and NEIL2 participate in oxidized base excision repair and protect cells from DNA damage. *NEIL1* (MIM:608844) and *NEIL2* (MIM:608933) variants may affect their protein functions, leading to altered cell-death and carcinogenesis. To date, only one reported study investigated the association between *NEIL1* and *NEIL2* polymorphisms and cancer risk.

Experimental Design—Genotype and haplotypes of the *NEIL1* NT_010194.16:g. 46434077G>T (rs7182283) and g.46438282C>G (rs4462560) and *NEIL2* NT_077531.3:g. 4102971C>G (rs804270) polymorphisms were determined for 872 patients with newly diagnosed SCCOOP and 1,044 cancer-free non-Hispanic white control subjects frequency matched by age and sex. Crude and adjusted odds ratios (OR) and 95% confidence intervals (95% CI) were calculated using multivariate logistic regression, and false-positive report probabilities were also calculated.

Results—We found no overall differences in the frequencies of alleles, genotypes, and haplotypes of *NEIL1* g.46434077G>T and *NEIL1* g.46438282C>G polymorphisms between cases and controls. However, the *NEIL2* g.4102971CC genotype was associated with a significantly increased risk of SCCOOP (adjusted OR, 1.30; 95% CI, 1.02–1.65); this increase in risk was the highest among current alcohol drinkers (adjusted OR, 1.87; 95% CI, 1.28–2.72), particularly in patients with oropharyngeal cancer (adjusted OR, 1.35; 95% CI, 1.04–1.76). The *NEIL2* g. 4102971CC genotype also was significantly associated with SCCOOP of advanced stages.

Conclusions—Polymorphisms of the *NEIL2* gene may be markers for risk and progression of SCCOOP, particularly in patients with oropharyngeal cancer. Larger studies are needed to confirm our findings.

Requests for reprints: Qingyi Wei, Department of Epidemiology, Unit 1365, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-3020; Fax: 713-563-0999; qwei@mdanderson.org.

Keywords

Genetic Susceptibility; Molecular Epidemiology; DNA Repair Capacity; Biomarkers; Base-Excision Repair

Introduction

Head and neck cancers are the sixth most common cancers, with a prevalence of more than 1.6 million cases worldwide (1). They account for approximately 2.8% of all malignancies in the United States (1,2). Despite slight decrease in its incidence rate (~ 4% since 1980) and a modest improvement in 5-year survival (from 54.4% to 59.4% over the last 20 years), these cancers continue to be a clinical challenge (1,2).

The squamous cell carcinomas of the oral cavity and oropharynx (SCCOOP) are an etiologically similar subgroup of head and neck cancers, and it is estimated that 34,360 new cases of SCCOOP and 7,550 deaths from it will have occurred in the United States alone in 2007 (2,3). Even with the use of modern therapeutic options such as surgical management, radiation treatment, and chemotherapeutic intervention, about 50% of all patients will ultimately die of this disease, especially for those patients who were diagnosed with advanced or relapsed disease, which is almost uniformly fatal (2).

The importance of an individual's lifestyle, particularly the use of tobacco and alcohol, has been well recognized in the etiology of SCCOOP. However, despite these risk factors, relatively few people actually develop these diseases; likewise, some individuals develop SCCOOP in the absence of such habits or other identifiable lifestyle and environmental factors. In this case, genetic susceptibility may play a role in the initiation and development of SCCOOP. For instance, genetically determined DNA repair capacity may contribute to the variation in susceptibility to head and neck cancers (4–6). Identifying genetic factors that modulate the risk of SCCOOP thus likely will help determine at-risk subgroups that can benefit from primary prevention programs.

DNA repair systems play an important role in the maintenance of genomic integrity and stability in response to environmental exposure, replication errors, and cumulative effects of aging. In humans, more than 100 genes are involved in the five major DNA repair pathways: direct reversal, basal excision repair (BER), nucleotide excision repair (NER), mismatch repair, and recombination repair (7). NER targets bulky, helix-distorting adducts, such as benzo(*a*)pyrene-guanine adduct, whereas BER removes smaller altered bases produced by oxidation, methylation, radiation, and other minor base modifications (8). In a series of association studies, we found that variants in genes involved in both BER and NER contributed to an individual's susceptibility to head and neck cancers (9–13).

In mammalian cells, the repair of DNA bases that have been damaged by reactive oxygen species (ROS) is initiated primarily by a series of DNA glycosylases, including two that belong to a class of DNA glycosylases homologous to the bacterial Fpg/Nei family, NEIL1 (nei endonuclease VIII–like 1) and NEIL2 (nei-like 2). These glycosylases initiate the first step in BER by cleaving bases damaged by ROS and introducing a DNA strand break via the associated lyase reaction (14–17). NEIL1 and NEIL2 protect cells from radiation-mediated cell death, and functional variants of *NEIL1* and *NEIL2* may affect the protein functions, leading to altered cell-death probability and carcinogenesis potential.

The human *NEIL1* gene is located on chromosome 15q23, and mutations in *NEIL1* have been shown to cause increased risk of developing primary gastric cancer in a Japanese

cohort (18). *NEIL1* has been resequenced by the National Institute of Environmental Health Sciences Environmental (NIEHS) Genome Project (EGP)[1], and at least 62 single nucleotide polymorphisms (SNPs)[2] have been reported; four of those are nonsynonymous SNPs (nsSNPs): p.Ser82Cys, p.Gly83Asp, p.Cys136Arg, and p.Asp252Asn, with very low minor allele frequencies (all = 0.01)[3]. Two of those polymorphic variants, p.Ser82Cys and p.Asp252Asn, showed near wild-type enzyme specificity and kinetics, whereas p.Gly83Asp was devoid of glycosylase activity and p.Cys136Arg may be glycosylase deficient (19). *NEIL2* is located on chromosome 8p23.1, and at least 250 SNPs in this gene are reported in the dbSNP database[4]. Four of these SNPs were confirmed as non-synonymous (nsSNPs) (p.Thr70Ser, p.Arg103Gln, p.Arg257Leu, and p.Pro304Thr) in the NIEHS EGP SNP database[5], and an additional one, P123T, was reported recently (20). All these nsSNPs have been found rare in both healthy control participants and colon cancer patients (20).

To date, no reports have been published about the possible association between variants of *NEIL1* and *NEIL2* genes and the risk of head and neck cancers. Because of the role of the *NEIL1* and *NEIL2* genes in regulating DNA repair and because decreased DNA repair capacity has been associated with increased risk of head and neck cancers, we hypothesized that *NEIL1* and *NEIL2* polymorphisms contribute to genetic susceptibility to SCCOOP. To test this hypothesis, we conducted a hospital-based case-control study to identify associations between SCCOOP and newly reported common (i.e., with a minor allele frequency of ≥ 0.05) functional SNPs of the *NEIL1* gene, NT_010194.16:g.46434077 (rs7182283) and NT_010194.16:g.46438282 (rs4462560), and the *NEIL2* gene, g.4102971 (rs804270).

Materials and Methods

Study subjects and data collection

The recruitment of our study participants has been previously described (21,22). Briefly, the study population consisted of 872 patients with SCCOOP and 1044 cancer-free control participants recruited from May 1, 1995, through September 30, 2006. All patients had newly diagnosed, untreated SCCOOP that was histologically confirmed at The University of Texas M. D. Anderson Cancer Center. Twenty four percent of all SCCOOP patients were eligible, and 90% of the eligible patients contacted chose to participate in this study. Only non-Hispanic white patients were included in the final analysis because genotype frequencies can vary between ethnic groups and relatively few minority patients were seen at M. D. Anderson. Among the 872 SCCOOP patients with primary tumors included in the analysis, 299 (34.3%) had cancers of the oral cavity and 573 (65.7%) cancers of the oropharynx. Patients with second oral cavity and oropharyngeal tumors; primary tumors of the hypopharynx, nasopharynx, or sinonasal tract; primary tumors outside the upper aerodigestive tract; cervical metastases of unknown origin; or histopathologic diagnoses other than SCCOOP were excluded from the analysis.

The regional lymph node involvement of SCCOOP was defined as N_0 to N_3 as follows (23): N_0, no regional node metastasis; N_1, metastasis in a single ipsilateral lymph node, ≤3 cm in the greatest dimension; N_2, metastasis in a single ipsilateral lymph node, >3 cm but <6 cm in the greatest dimension; or in multiple ipsilateral lymph nodes, none ≥6 cm in the greatest dimension; or in any bilateral or contralateral lymph node, <6 cm in the greatest dimension;

[1]http://egp.gs.washington.edu/data/NEIL1/.
[2]http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=79661&chooseRs=all.
[3]http://egp.gs.washington.edu/data/NEIL1/NEIL1.csnps.txt.
[4]http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=252969&chooseRs=all.
[5]http://egp.gs.washington.edu/data/NEIL2/NEIL2.csnps.txt.

N_3, metastasis in any lymph node, ≥6 cm in the greatest dimension. The extent of the primary SCCOOP was defined as T_1 to T_4 as follows: T_1, tumor ≤2 cm at the greatest dimension; T_2, tumor >2 cm but <4 cm in the greatest dimension; T_3, tumor ≥4 cm in the greatest dimension; T_4, tumor invading adjacent structures.

Cancer-free control participants were recruited from persons who were not hospital patients or seeking health care but who had accompanied the case patients visiting the clinics; controls could not be genetically related to any cases or controls already selected. We first surveyed potential control participants at the clinics by using a short questionnaire to determine their willingness to participate in our research studies and to obtain information about demographic features, smoking and alcohol drinking status (current, former, or never), and personal history of cancer. Our control participants were required to have no previous history of cancer and were not under medical care or receiving treatment for any known disease. They were frequency matched to the case patients by age (± 5 years) and sex. Among the visitors to our institution who were screened for possible participation as controls, 73% were eligible for and were therefore offered participation. Of these eligible potential controls, 85% agreed to and ultimately did participate. We interviewed each enrolled control subject to obtain data about their personal history of exposure to known etiologic factors in SCCOOP, such as tobacco smoking and alcohol use. After signing the informed consent form, each participant donated 30 mL of blood, of which 1 mL was used for genomic DNA extraction. The protocol for this research study was approved by the M. D. Anderson Cancer Center Institutional Review Board.

Genotyping

We extracted genomic DNA from the buffy-coat fraction of the blood samples by using a blood DNA mini kit (Qiagen Inc., Valencia, CA) according to the manufacturer's instructions. The selected *NEIL1* and *NEIL2* polymorphisms were determined using the polymerase chain reaction (PCR)–restriction fragment length polymorphism method, as previously described (21). We performed the PCRs with a PTC-200 DNA engine (Peltier thermal cycler, M J Research, Inc., Waltham, MA) in 10 μL of PCR mixture. This PCR mixture included approximately 20 ng of genomic DNA, 0.1 mM deoxynucleotide triphosphate, and 1× PCR buffer (50 mM KCl, 10 mM Tris HCl, and 0.1% Triton X-100); 1.5 mM $MgCl_2$, 0.5 units of Taq polymerase (Denville Scientific Inc.; Metuchen, NJ); and 2 pmol of each primer. The following primers were used to amplify the target fragments containing these four polymorphisms (mismatch bases are underlined): 5′-CAACCTCCTGATTAACTGGAACCACA -3′ (forward) and 5′-TCACTTCAGCCCAGGAGACCAG -3′ (reverse) for *NEIL1* NT_010194.16:g.46434077G>T (rs7182283); 5′-GTCTCTTCACTGGCTTTTGGGG -3′ (forward) and 5′-TCCCAGGTATTTGGTGGGTAGG -3′ (reverse) for *NEIL1* NT_010194.16:g.46438282C>G (rs4462560); and 5′-ACCCCCCACCTCGGGCACTCGG -3′ (forward) and 5′-AAGATGCCGCGCCCACCCGC -3′ (reverse) for *NEIL2* NT_077531.3:g.4102971C>G (rs804270). There were 119-bp, 118-bp, and 131-bp amplified PCR products for the *NEIL1* g.46434077, *NEIL1* g.46438282, and *NEIL2* g.4102971 polymorphisms, respectively. The *Nla*III, *Hae*III, and *Sac*II restriction enzymes (New England Biolabs, Beverly, MA) were used to distinguish the *NEIL1* g.46434077, *NEIL1* g.46438282, and *NEIL2* g.4102971 polymorphisms, respectively, which resulted in 91-bp and 28-bp fragments in the presence of the *NEIL1* g.46434077T allele; 96-bp and 22-bp fragments in the presence of the *NEIL1* g.46438282C allele; and 112-bp and 19-bp fragments in the presence of the *NEIL2* g.4102971C allele. More than 10% of the samples were randomly selected for confirmation, and the results were 100% concordant.

Statistical analysis

The χ^2 test was used to evaluate differences between cases and controls in the frequency distributions of selected demographic variables, smoking status, alcohol use, and each allele and genotype of the three *NEIL1* and *NEIL2* polymorphisms. Unconditional univariate and multivariate logistic regression analyses were performed to obtain the crude and adjusted odds ratios (ORs) and their 95% confidence intervals (CIs) for cancer risk, assuming a genetic recessive model. Multivariate adjustment was conditional on the effect of age, sex, smoking status, and alcohol use. Subjects who had smoked > 100 cigarettes in their lifetimes were categorized as ever smokers, and others were never smokers. Ever smokers who had quit smoking > 1 year previously (prior to disease diagnosis for the cases and before the time of questionnaire administered for the controls) were categorized as former smokers, and the other smokers were categorized as current smokers. Similarly, subjects who had drunk alcoholic beverages at least once a week for > 1 year previously were categorized as ever drinkers, and others were never drinkers. Ever drinkers who had quit drinking > 1 year previously were categorized as former drinkers, and the other drinkers were categorized as current drinkers. We further stratified the genotype data by subgroups of age, sex, smoking, alcohol drinking, and tumor site and assessed any trend in risk in multivariate logistic regression models. Hardy–Weinberg equilibrium (HWE) of the *NEIL1* and *NEIL2* genotypes was tested by performing a goodness-of-fit χ^2-test. The linkage disequilibrium D′ value, logarithm of odds (LOD) score, and r^2 value were computed for two-locus models. We estimated *NEIL1* haplotypes using unphased genotypes. The haplotype of the highest frequency was used as the reference group to calculate odds ratios for haplotypes associated with SCCOOP. Logistic regression was used to evaluate the association between the SCCOOP and *NEIL1* haplotypes. Haplotypes were reconstructed using PHASE version 2 software (24)[6]

We used the false-positive report probability (FPRP) to test for false-positive associations (25). For all significant genetic effects observed in our study, we calculated FPRP with prior probabilities of 0.0001, 0.001, 0.01, 0.1 and 0.25. The OR was set close to the observed value obtained in our study, and a probability of < 0.2 was considered noteworthy. We analyzed our data by using SAS software (Version 8e; SAS Institute Inc., Cary, NC); $P < 0.05$ was considered statistically significant, and all statistical tests were two sided,.

Results

Characteristics of non-genetic risk factors for SCCOOP in the study population

The frequency distributions of selected characteristics of the cases and controls are presented in Table 1. The cases and controls were well matched for age and sex, with a mean age of 56.5 years for the cases (± 11.2 years; range, 18–85 years) and 56.4 years for the controls (± 11.0 years; range, 20–85 years) ($P = 0.975$); there was also no difference in the distribution of age groups (≤ 50, 51–64, and ≥ 65) nor in sex groups between the case and control groups, with men constituting 77.2% of patients and 76.8% of controls ($P = 0.853$). However, there were more current smokers (33.6%) and current drinkers (51.7%) among the cases than among the controls (15.3% and 40.7%, respectively), and this differences were statistically significant ($P < 0.001$). There was an increased risk of SCCOOP in former and current smokers (OR, 1.62; 95% CI, 1.31–2.01) and OR, 3.55; 95% CI, 2.78–4.53, respectively) and in former or current alcohol drinkers (OR, 2.06; 95% CI, 1.59–2.68, and OR, 2.01; 95% CI, 1.64–2.47, respectively) compared with non-smokers and non-drinkers, respectively (Table 1). Therefore, we made further adjustment for these covariates in the multivariate logistic regression analysis.

[6]http://stephenslab.uchicago.edu/software.html.

Allele and genotype distributions of the NEIL1 and NEIL2 polymorphisms among the cases and controls

As shown in Table 2, the frequencies of minor alleles of *NEIL1* g.46434077T and g.46438282G and *NEIL2* g.4102971C were 47.2%, 31.5% and 44.4%, respectively, for the cases and 46.4%, 31.8% and 43.0%, respectively, for the controls, but none of the differences between the cases and controls was statistically significant ($P = 0.611$, 0.809 and 0.359, respectively). All distributions of the frequencies of observed genotypes in the controls were consistent with those obtained from the Hardy-Weinberg equilibrium model ($P > 0.05$). The *NEIL1* g.46434077TT genotype was more frequent in the cases (22.3%) than in the controls (20.6%) but was not associated with risk of SCCOOP compared with either the GG genotype (OR, 0.99; 95% CI, 0.76–1.29) or the combined GG+GT genotypes (OR, 1.03; 95% CI, 0.82–1.29). The *NEIL1* g.46438282GG genotype was less frequent in the cases (8.9%) than in the controls (9.5%), and similarly was not associated with risk of SCCOOP compared with either the CC genotype (OR, 0.99; 95% CI, 0.71–1.39) or the combined CC+CG genotypes (OR, 0.97; 95% CI, 0.70–1.34). The *NEIL2* g.4102971CC genotype was more frequent in the cases (20.7%) than in the controls (17.1%), and although this genotype was not associated with risk of SCCOOP compared with the *NEIL2* g.4102971GG genotype (adjusted OR, 1.18; 95% CI, 0.90–1.56), it was associated with a significant elevation in risk when compared with the combined GG+CG genotypes (OR, 1.30; 95% CI, 1.02–1.65; $P < 0.05$, assuming a recessive genetic model).

Association between the combined genotypes and Haplotype of the NEIL1 gene and the risk of SCCOOP

Further analysis suggested that there was a linkage disequilibrium between the two *NEIL1* polymorphisms ($D' = 1.0$, $r^2 = 0.304$, and $P < 0.001$). Since haplotype preserves the joint linkage disequilibrium structure, we reconstructed haplotypes using the unphased genotypes based on the observed *NEIL1* g.46434077 and g.46438282 genotypes to further assess their association with risk of SCCOOP. There were four haplotypes: TC, GG, GC and TG; haplotype TC was the most common (occurring in 43.02% of the cases and 43.44% of the controls) and haplotype TG was the least frequent, with a frequency < 5% in both cases and controls. The frequency distributions of these four haplotypes were comparable between cases and controls (χ_2 test: $P = 0.87$). Therefore, neither the genotypes nor haplotypes of *NEIL1* g.46434077 and *NEIL1* g.46438282 were associated with risk of SCCOOP (data not sown).

Stratified analysis of the associations between NEIL1 and NEIL2 polymorphisms and risk of SCCOOP

We further stratified the data by age, sex, smoking status, alcohol drinking status, and cancer site. To facilitate the analysis, we evaluated the risk of SCCOOP by estimating the ORs associated with the combined *NEIL1* g.46434077 (GG+GT), *NEIL1* g.46438282 (CC+CG), and *NEIL2* g.4102971 (GG+CG) genotypes compared with their homozygous genotypes (*NEIL1* g.46434077TT, *NEIL1* g.46438282GG, and *NEIL2* g.4102971CC, respectively), with adjustment for the aforementioned variables. As shown in Table 3, although none of the genotypes of the *NIEL1* gene was associated with cancer risk in all subgroups examined, the OR was virtually always elevated (except for former alcohol drinkers), sometimes significantly so, for the CC genotype of *NEIL2* g.4102971 compared with the combined GG+CG genotypes. In former smokers and current drinkers, the OR the CC genotype of *NEIL2* g.4102971 were 1.51 (95% CI, 1.03–2.23; $P = 0.082$) and 1.87 (95% CI, 1.28–2.72; $P = 0.003$), respectively, compared with either the GG or CG genotype, and this risk was more evident for oropharynx (OR, 1.35; 95% CI, 1.04–1.76; $P = 0.02$) than for oral cavity (OR, 1.19; 95% CI, 0.83–1.69; $P = 0.545$), although this difference was not statistically significant (data not shown).

Association between the NEIL2 g.4102971 polymorphisms and progression of SCCOOP

Because only the *NEIL2* g.4102971 polymorphism was associated with the risk of SCCOOP, we further evaluated its association with progression of SCCOOP. We found a statistically significant association between the *NEIL2* g.4102971CC genotype and advanced stages of primary SCCOOP tumor (T) as well as regional lymph node metastasis at diagnosis (N). The genotype frequency distributions were statistically significantly different between cases and controls for tumor stages T_2 and T_3 and for node stages N_2 and N_3. Using the 1044 controls as the reference group, the adjusted ORs for the CC genotype of *NEIL2* g. 4102971 were calculated for each stratum by T or N stage. ORs for the CC genotype were significantly increased as the stages increased with the highest risk for T_3 and N_3 stage after adjustment for other covariates. For example, the OR for the genotype CC of *NEIL2* g. 4102971 in SCCOOP patients with N_3 was 3.06 (95% CI, 1.49–6.26). When we grouped SCCOOP patients into two groups according to their T and N status (i.e., a group with less-advanced disease [N_0 or N_1 and T_1 or T_2] and a group with more-advanced disease [N2, N3, T3, or T4], the OR for the genotype CC vs. other genotypes was 1.39 (95% CI, 1.07–1.81) in SCCOOP patients with more-advanced disease, compared with those patients with less-advanced disease, after adjusting for other covariates (Table 4).

Because most of the significant findings were in the subgroup analysis, we calculated the FPRP value for all the significant associations. As shown in Table 5, when the assumption of prior probability was 0.1, the *NEIL2* g.4102971CC genotype was still associated with an elevated risk of SCCOOP in current alcohol drinkers and in those with tumor stage N_3, with FPRP = 0.05 and 0.1, respectively. However, if the prior probability was set to 0.01, all significant associations we found in this study became false-positive findings.

Discussion

In this study, we investigated the associations between three common, potentially functional polymorphisms of *NEIL1* and *NEIL2* genes and the risk of SCCOOP in a hospital-based case-control study. When we evaluated each polymorphism separately, only the *NEIL2* g. 4102971 polymorphism, but not *NEIL1* g.46434077 or g.46438282, was associated with risk of SCCOOP. In addition, the *NEIL2* g.4102971 polymorphism appeared to be associated with more advanced SCCOOP, particularly for oropharyngeal cancer. Given the role of the *NEIL* genes in the BER pathway, it is biologically plausible that functional *NEIL2* polymorphisms may modulate the risk and/or progression of cancer.

Several groups have reported that *NEIL1* and *NEIL2* sequence variants are associated with the risk of such diseases as colorectal cancer, gastric cancer, and metabolic syndrome as well as brain ontogeny in different ethnic groups (20,26,27). One study suggested that mutations in *NEIL1* may reduce the gene expression and protein activities in a subset of gastric cancers (18), but another study reported that some sequence variations in both *NEIL1* and *NEIL2* were reportedly not associated with the risk of colorectal cancer (20). However, no previous studies investigated associations between *NEIL1* and *NEIL2* polymorphisms and the risk of SCCOOP. Our current large case-control study support the notion that selected variants in *NEIL2* may contribute to the etiology of SCCOOP.

Several lines of evidence of molecular mechanisms support our findings. Oxidized DNA base lesions such as thymine glycol and 8-hydroxyguanine are often toxic, mutagenic or even carcinogenic (12). In mammalian cells, the repair of DNA bases that have been damaged by ROS is primarily initiated by a series of DNA glycosylases that include OGG1, NTH1, NEIL1 and NEIL2. It has been shown that *in vitro* translated mouse or human NEIL1 can remove thymine glycol and 5-hydroxyuracil much more efficiently than 8-oxoG in double- and single-strand DNA (28,29). With the loss of the NEIL1 functions, *NEIL1*

knockout mice develop metabolic syndrome, manifesting as severe obesity, dyslipidemia, and fatty liver disease (26). The genetic variants of *NEIL1* selected in the present study may not have been sufficient to cause severe loss of NEIL1 functions; alternatively, our study may not have had enough power to detect small differences in terms of cancer risk, if any existed. NEIL2 primarily functions to excise oxidative products of cytosine, with its greatest activity for 5-hydroxyuracil, but it shows negligible or undetectable activity for 8-oxoG (14). Our finding of an association between the selected *NEIL2* variant and risk of SCCOOP could be due to a function loss; alternatively, this variant may be in linkage disequilibrium with other causal variants. These hypotheses should be tested in future mechanistic studies.

In the present study, we observed an increased risk associated with the variant *NIEL2* genotype in former smokers but not current smokers, and this finding is likely to be a chance finding. Likewise, we also observed a significantly increased risk of SCCOOP among current drinkers of alcohol, suggesting that a gene-environment interaction may be involved in the development of SCCOOP; however, our study did not have enough statistical power to confirm any such gene-environment interaction. Ethanol reportedly induces oxidative stress via metabolic activation, leading to oxidative DNA damage and a decrease in hepatic antioxidant defense (27,30), which may explain the putative role of the *NEIL2* g.4102971 polymorphism in alcohol-induced SCCOOP. However, this finding may also have occurred by chance, owing to the small number of observations in our stratified analysis.

More interesting is our finding that *NEIL2* g.4102971CC genotype was significantly associated with more-advanced SCCOOP, suggesting that the *NEIL2* g.4102971CC genotype may be associated with SCCOOP progression, although it is also possible that our results may be due to selection bias commonly occurring in hospital-based case-control studies. However, we can speculate that because NEIL2 is involved in the repair of oxidative damage to DNA either accumulated in the aging process or as a result of fast-growing tumors due to their enhanced metabolic activities (31), an altered function of NEIL2 due to the variant may lead to excessive oxidative damage to DNA, causing additional mutations or a mutator phenotype with genomic instability that may promote tumor progression (32). This hypothesis needs to be tested in additional mechanistic studies.

It is difficult to compare our genotyping data with those from other studies, because few studies on *NEIL1* and *NEIL2* polymorphisms have been published. Additionally, because other less common functional variants of these two genes were not assayed in the present study, our finding must be confirmed in studies that enroll larger numbers of patients with SCCOOP with genotyping data based on dense gene maps, such as the HapMap database (33).

In conclusion, we found that the *NEIL2* g.4102971CC genotype was associated with a significantly increased risk of SCCOOP, particularly advanced SCCOOP or oropharygeal cancer, compared with other genotypes. However, the FPRP values for all the significant findings in our study were greater that 0.2 when the prior probability was set at 0.01, suggesting that these findings could all be false positives. Therefore, our findings should validated in future population-based studies that include larger numbers of patients with oropharyngeal cancer, more detailed data on environmental exposure, more SNPs in more genes in the same biologic pathway, and survival data. Because the majority of patients did not have enough follow-up time or death events yet, we will evaluate the role of *NEIL2* g.4102971CC genotype on disease prognosis in this patient cohort in the future follow-up study.

Acknowledgments

Grant support: NIH grants ES11740 (Q. Wei), CA100264 (Q. Wei), and CA16672 (The University of Texas M. D. Anderson Cancer Center).

We thank Margaret Lung, Kathryn Patterson, and Leanel Fairly for their assistance in recruiting the subjects; Yawei Qiao, Jianzhong He, and Kejing Xu for their laboratory assistance; and Kathryn Carnes and Chris Hengst for scientific editing.

References

1. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. N Engl J Med 2001;345:1890–900. [PubMed: 11756581]
2. Ries, L.; Krapcho, M.; Mariotto Aea. SEER Cancer Statistics Review, 1975–2003. National Cancer Institute; Bethesda, MD: 2006.
3. ACS. American Cancer Society Cancer Facts & Figures 2007. Atlanta, GA: 2007.
4. Cheng L, Eicher SA, Guo Z, Hong WK, Spitz MR, Wei Q. Reduced DNA repair capacity in head and neck cancer patients. Cancer Epidemiol Biomarkers Prev 1998;7:465–8. [PubMed: 9641488]
5. Neumann AS, Sturgis EM, Wei Q. Nucleotide excision repair as a marker for susceptibility to tobacco-related cancers: a review of molecular epidemiological studies. Mol Carcinog 2005;42:65–92. [PubMed: 15682379]
6. Xiong P, Hu Z, Li C, et al. 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. Int J Cancer. 2007
7. Sancar A. DNA excision repair. Annu Rev Biochem 1996;65:43–81. [PubMed: 8811174]
8. Yu Z, Chen J, Ford BN, Brackley ME, Glickman BW. Human DNA repair systems: an overview. Environ Mol Mutagen 1999;33:3–20. [PubMed: 10037319]
9. Sturgis EM, Castillo EJ, Li L, et al. Polymorphisms of DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis 1999;20:2125–9. [PubMed: 10545415]
10. Sturgis EM, Zheng R, Li L, et al. XPD/ERCC2 polymorphisms and risk of head and neck cancer: a case-control analysis. Carcinogenesis 2000;21:2219–23. [PubMed: 11133811]
11. Shen H, Sturgis EM, Dahlstrom KR, Zheng Y, Spitz MR, Wei Q. A variant of the DNA repair gene XRCC3 and risk of squamous cell carcinoma of the head and neck: a case-control analysis. Int J Cancer 2002;99:869–72. [PubMed: 12115490]
12. Li C, Hu Z, Lu J, et al. Genetic polymorphisms in DNA base-excision repair genes ADPRT, XRCC1, and APE1 and the risk of squamous cell carcinoma of the head and neck. Cancer 2007;110:867–75. [PubMed: 17614107]
13. An J, Liu Z, Hu Z, et al. Potentially functional single nucleotide polymorphisms in the core nucleotide excision repair genes and risk of squamous cell carcinoma of the head and neck. Cancer Epidemiol Biomarkers Prev 2007;16:1633–8. [PubMed: 17684138]
14. Hazra TK, Kow YW, Hatahet Z, et al. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J Biol Chem 2002;277:30417–20. [PubMed: 12097317]
15. Hazra TK, Mitra S. Purification and characterization of NEIL1 and NEIL2, members of a distinct family of mammalian DNA glycosylases for repair of oxidized bases. Methods Enzymol 2006;408:33–48. [PubMed: 16793361]
16. Bandaru V, Sunkara S, Wallace SS, Bond JP. A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII. DNA Repair (Amst) 2002;1:517–29. [PubMed: 12509226]
17. Das A, Wiederhold L, Leppard JB, et al. NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells. DNA Repair (Amst) 2006;5:1439–48. [PubMed: 16982218]
18. Shinmura K, Tao H, Goto M, et al. Inactivating mutations of the human base excision repair gene NEIL1 in gastric cancer. Carcinogenesis 2004;25:2311–7. [PubMed: 15319300]

19. Roy LM, Jaruga P, Wood TG, McCullough AK, Dizdaroglu M, Lloyd RS. Human polymorphic variants of the NEIL1 DNA glycosylase. J Biol Chem 2007;282:15790–8. [PubMed: 17389588]
20. Broderick P, Bagratuni T, Vijayakrishnan J, Lubbe S, Chandler I, Houlston RS. Evaluation of NTHL1, NEIL1, NEIL2, MPG, TDG, UNG and SMUG1 genes in familial colorectal cancer predisposition. BMC Cancer 2006;6:243. [PubMed: 17029639]
21. Li G, Sturgis EM, Wang LE, et al. Association between the V109G polymorphism of the p27 gene and the risk and progression of oral squamous cell carcinoma. Clin Cancer Res 2004;10:3996–4002. [PubMed: 15217930]
22. 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 Res 2007;67:5628–34. [PubMed: 17575128]
23. Cancer AJCo. Manual for Staging of Cancer. 6. Springer, Philadelphia: JB Lippincott; 2002.
24. Stephens JC, Schneider JA, Tanguay DA, et al. Haplotype variation and linkage disequilibrium in 313 human genes. Science 2001;293:489–93. [PubMed: 11452081]
25. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42. [PubMed: 15026468]
26. Vartanian V, Lowell B, Minko IG, et al. The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase. Proc Natl Acad Sci U S A 2006;103:1864–9. [PubMed: 16446448]
27. Dey A, Cederbaum AI. Alcohol and oxidative liver injury. Hepatology 2006;43:S63–74. [PubMed: 16447273]
28. Takao M, Kanno S, Kobayashi K, et al. A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. J Biol Chem 2002;277:42205–13. [PubMed: 12200441]
29. Dou H, Mitra S, Hazra TK. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. J Biol Chem 2003;278:49679–84. [PubMed: 14522990]
30. Seitz HK, Stickel F. Risk factors and mechanisms of hepatocarcinogenesis with special emphasis on alcohol and oxidative stress. Biol Chem 2006;387:349–60. [PubMed: 16606331]
31. Loft S, Poulsen HE. Cancer risk and oxidative DNA damage in man. Journal of molecular medicine (Berlin, Germany) 1996;74:297–312.
32. Charames GS, Bapat B. Genomic instability and cancer. Current molecular medicine 2003;3:589–96. [PubMed: 14601634]
33. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature 2005;437:1299–320. [PubMed: 16255080]

Table 1

Characteristics and odds ratios of covariate variables in the study

Variable	Cases (n = 872) *n* (%)	Controls (n = 1,044) *n* (%)	Odds Ratio*	*P*†
Mean age (± SD)‡	56.5 (±11.2)	56.4 (±11.0)		0.975
Age				
≤ 50	245 (28.1)	303 (29.0)	Reference	0.898
51–64	410 (47.0)	487 (46.7)	1.04 (0.84–1.29)	
≥ 65	217 (24.9)	254 (24.3)	1.06 (0.83–1.35)	
Gender				
Female	199 (22.8)	242 (23.2)	Reference	0.853
Male	673 (77.2)	802 (76.8)	1.02 (0.82–1.26)	
Smoking status				
Never	259 (29.7)	502 (48.1)	Reference	<0.001
Former	320 (36.7)	382 (36.6)	1.62 (1.31–2.01)	
Current	293 (33.6)	160 (15.3)	3.55 (2.78–4.53)	
Alcohol use				
Never	238 (27.3)	451 (43.2)	Reference	<0.001
Former	183 (21.0)	168 (16.1)	2.06 (1.59–2.68)	
Current	451 (51.7)	425 (40.7)	2.01 (1.64–2.47)	
Tumor site				
Oral cavity	299 (34.3)			
Oropharynx	573 (65.7)			

*Odds ratio was calculated for categorical variables of age, gender, smoking, and alcohol use.

†*P* value is from the χ^2 *test* for categorical variables. The *P* value for mean age was from the Wilcoxon rank test.

‡Mean age (± SD) was the mean and standard deviation of variable age.

Table 2

Genotype frequencies and odds ratios of *NEIL1* g.46434077 and g.46438282 and *NEIL2* g.4102971 variants among SCCOOP cases and controls in a non-Hispanic white population

SNP	Cases *n* (%)	Controls *n* (%)	Crude Odds ratio	Adjusted Odds ratio*	*P*†
NEIL1 g.46434077G>T‡					
GG	243 (28.0)	291 (27.9)	Reference	Reference	0.624
GT	432 (49.7)	537 (51.5)	0.96 (0.78–1.19)	0.94 (0.76–1.17)	
TT	194 (22.3)	215 (20.6)	1.08 (0.83–1.4)	0.99 (0.76–1.29)	
GG+GT	675 (77.7)	828 (79.4)	Reference	Reference	0.364
TT	194 (22.3)	215 (20.6)	1.11 (0.89–1.38)	1.03 (0.82–1.29)	
T allele	820 (47.2)	967 (46.4)			0.611
NEIL1 g.46438282C>G§					
CC	392 (46.0)	477 (45.8)	Reference	Reference	0.902
CG	385 (45.1)	465 (44.7)	1.01 (0.83–1.22)	1.05 (0.87–1.28)	
GG	76 (8.9)	99 (9.5)	0.93 (0.67–1.3)	0.99 (0.71–1.39)	
CC+CG	777 (91.1)	942 (90.5)	Reference	Reference	0.656
GG	76 (8.9)	99 (9.5)	0.93 (0.68–1.27)	0.97 (0.70–1.34)	
G allele	537 (31.5)	663 (31.8)			0.809
NEIL2 g.4102971G>C‖					
GG	273 (31.9)	322 (31.1)	Reference	Reference	0.069
CG	405 (47.4)	538 (51.9)	0.89 (0.72–1.09)	0.86 (0.69–1.06)	
CC	177 (20.7)	177 (17.1)	1.18 (0.91–1.54)	1.18 (0.90–1.56)	
GG+CG	678 (79.3)	860 (82.9)	Reference	Reference	0.044
CC	177 (20.7)	177 (17.1)	1.27 (1.01–1.6)	1.30 (1.02–1.65)	
C allele	759 (44.4)	892 (43.0)			0.359

*Adjusted odds ratio: adjusted for age, gender, smoking, and alcohol use.

†P value is from the χ^2 *test* for the frequency distribution of genotypes in each SNP.

‡*NEIL1* g.46434077G>T: missing genotype information in 3 cases and 1 control.

§*NEIL1* g.46438282C>G: missing genotype information in 19 cases and 3 controls.

||*NEIL2* g.41029710G>C: missing genotype information in 17 cases and 7 controls.

Table 3

Odds ratios of SCCOOP associated with genotypes in the stratified analysis

Stratified variables	*n* (case/control)*	*NEIL1* g.46434077G>T		*P*‡	*n* (case/control)§	*NEIL1* g.46438282C>G		*P*‡	*n* (case/control)‖	*NEIL2* g.4102971G>C		*P*‡
		Percentage (case/control)	Adjusted OR (95% CI)†			Percentage (case/control)	Adjusted OR (95% CI)†			Percentage (case/control)	Adjusted OR (95% CI)†	
		TT	TT vs. GG + GT			GG	GG vs. CC + CG			CC	CC vs. CG + GG	
Age												
≤ 50	243/303	21.0/20.5	0.98 (0.64–1.50)	0.88	237/303	10.5/7.6	1.47 (0.80–2.68)	0.231	236/301	19.9/16.9	1.25 (0.80–1.95)	0.376
51 to 64	409/487	24.0/20.5	1.06 (0.75–1.48)	0.218	404/485	9.7/11.1	0.89 (0.56–1.42)	0.473	404/484	20.8/18.2	1.27 (0.90–1.81)	0.327
≥ 65	217/253	20.7/20.9	0.92 (0.58–1.46)	0.955	212/253	5.7/8.7	0.63 (0.30–1.32)	0.211	215/252	21.4/15.1	1.40 (0.85–2.28)	0.077
Gender												
Male	671/801	23.4/20.5	1.11 (0.86–1.43)	0.176	662/800	8.6/9.1	1.00 (0.69–1.45)	0.731	659/797	21.1/17.8	1.27 (0.97–1.66)	0.115
Female	198/242	18.7/21.1	0.76 (0.46–1.25)	0.533	191/241	9.9/10.8	0.86 (0.44–1.66)	0.776	196/240	19.4/14.6	1.49 (0.87–2.54)	0.181
Smoking status												
Never	258/502	18.6/18.7	0.98 (0.67–1.44)	0.968	252/502	7.5/10.6	0.68 (0.39–1.18)	0.184	251/499	19.5/17.6	1.13 (0.76–1.66)	0.528
Former	320/381	20.9/22.6	0.91 (0.63–1.31)	0.602	315/379	10.5/8.4	1.35 (0.81–2.27)	0.36	317/380	21.8/16.6	1.51 (1.03–2.23)	0.082
Current	291/160	27.1/21.9	1.29 (0.80–2.06)	0.218	286/160	8.4/8.8	0.96 (0.46–1.97)	0.897	287/158	20.6/16.5	1.41 (0.83–2.39)	0.292
Alcohol use												
Never	237/451	21.5/20.6	1.05 (0.72–1.55)	0.783	234/450	9.8/11.1	0.87 (0.52–1.47)	0.606	235/449	20.9/19.6	1.10 (0.74–1.63)	0.698
Former	183/168	19.7/16.1	1.26 (0.72–2.20)	0.38	179/168	5.6/10.1	0.51 (0.23–1.16)	0.115	181/166	18.8/18.7	0.97 (0.56–1.69)	0.979
Current	449/424	23.8/22.4	0.98 (0.70–1.37)	0.618	440/423	9.8/7.6	1.43 (0.87–2.38)	0.25	439/422	21.4/13.7	1.87 (1.28–2.72)	0.003
Tumor site												
Oral cavity	297/1043	22.6/20.6	1.04 (0.75–1.44)	0.468	294/1041	7.8/9.5	0.83 (0.50–1.36)	0.375	296/1037	18.6/17.1	1.19 (0.83–1.69)	0.545
Oropharynx	572/1043	22.2/20.6	1.03 (0.80–1.33)	0.455	559/1041	9.5/9.5	1.04 (0.73–1.49)	0.985	559/1037	21.8/17.1	1.35 (1.04–1.76)	0.020

The result is bolded if the 95% confidence interval does not include 1 or $P < 0.05$.

*No. (cases/controls): 3 cases and 1 control had missing values for *NEIL1* g.46434077G>T.

†Adjusted OR, odds ratios were adjusted for all covariates (age, gender, smoking status, and alcohol use), excluding the stratified variable.

‡P value was from the χ^2 *test* statistics from comparisons of genotype frequency in cases and controls.

§No. (cases/controls): 19 cases and 3 controls had missing values for *NEIL1* g.46438282C>G.

‖No. (cases/controls) 17 cases and 7 controls had missing values for *NEIL2* g.4102971G>C.

Table 4

Associations between the combined genotypes of the *NEIL2* g.4102971 polymorphism and progression of SCCOOP*

Stratified variables*	*NEIL2* g.4102971G>C			Crude OR	Adjusted OR †	*P*‡
	n (case/control)	CC	GG+CG			
Primary tumor (T)						
T_1	216/1037	16.2/17.1	83.8/82.9	0.94 (0.63–1.40)	0.98 (0.65–1.46)	0.758
T_2	315/1037	22.5/17.1	77.5/82.9	1.41 (1.04–1.93)	1.48 (1.07–2.03)	0.028
T_3	159/1037	24.5/17.1	75.5/82.9	1.58 (1.06–2.35)	1.75 (1.16–2.64)	0.023
T_4	165/1037	19.4/17.1	80.6/82.9	1.17 (0.77–1.78)	1.19 (0.76–1.87)	0.464
Lymph node metastasis						
N_0	277/1037	19.5/17.1	80.5/82.9	1.18 (0.84–1.65)	1.24 (0.87–1.77)	0.346
N_1	127/1037	15.7/17.1	84.3/82.9	0.91 (0.55–1.50)	1.01 (0.60–1.69)	0.708
N_2	413/1037	21.8/17.1	78.2/82.9	1.35 (1.02–1.80)	1.40 (1.05–1.88)	0.036
N_3	38/1037	34.2/17.1	65.8/82.9	2.53 (1.27–5.03)	3.06 (1.49–6.26)	0.007
Combine N and T						
$N_{0,1}$ and $T_{1,2}$	260/1037	18.8/17.1	81.2/82.9	1.13 (0.79–1.60)	1.19 (0.83–1.72)	0.500
$N_{2,3}$ or $T_{3,4}$	595/1037	21.5/17.1	78.5/82.9	1.33 (1.03–1.72)	1.39 (1.07–1.81)	0.027

Note: The result is bolded if the 95% confidence interval does not include 1 or $P < 0.05$.

*Stratified variables: T: The extent of the primary SCCOOP. T_1: tumor 2 cm at the greatest dimension; T_2 to T_4: increasing greatest dimensions. N: regional lymph node involvement. N_0: no regional lymph nodes involved; N_1 to N_3: increasing involvement of regional lymph nodes.

†Adjusted OR: odds ratios were adjusted for age, gender, smoking status, and alcohol use.

‡*P* value from the χ^2 test of different frequencies of *NEIL2* g.4102971 genotypes in cases and controls.

Table 5

FPRP for significant associations of *NEIL2* g.4102971 and SCCOOP

Subgroup	Positive OR95%CI*	*P*†	OR‡	Statistical power§	Prior probability				
					0.25	0.1	0.01	0.001	0.0001
All participants	1.27 (1.01–1.60)	0.044	1.2	0.69	0.16	0.36	0.86	0.98	1.00
Current alcohol users	1.71 (1.19–2.45)	0.003	1.7	0.51	0.02	0.05	0.38	0.86	0.98
Oropharynx cancers	1.36 (1.05–1.76)	0.02	1.3	0.63	0.09	0.22	0.76	0.97	1.00
T_2	1.41 (1.04–1.93)	0.028	1.4	0.53	0.14	0.32	0.84	0.98	1.00
T_3	1.58 (1.06–2.35)	0.023	1.5	0.6	0.10	0.25	0.79	0.97	1.00
N_2	1.35 (1.02–1.80)	0.036	1.3	0.61	0.15	0.35	0.85	0.98	1.00
N_3	2.53 (1.27–5.03)	0.007	2.5	0.51	0.04	0.10	0.56	0.93	0.99
$N_{2,3}$ or $T_{3,4}$	1.33 (1.03–1.72)	0.027	1.3	0.58	0.12	0.29	0.82	0.98	1.00

*OR 95% CI calculated from the study sample by univariate logistic regression analysis.

†*P* value calculated by the χ^2 test in comparisons of genotype distributions in cases and controls.

‡OR was the odds ratio closest to the observed odds ratio in each stratified subgroup used to calculate FPRP.

§Statistical power was calculated using the number of observations in the subgroup and the OR and *P* values in this table.