Molecular epidemiology of DNA repair gene polymorphisms and head and neck cancer

Meilin Wang\textsuperscript{a,b}, Haiyan Chu\textsuperscript{a,b}, Zhengdong Zhang\textsuperscript{a,b,\textcopyright}, Qingyi Wei\textsuperscript{a,\textcopyright}

\textsuperscript{a}Department of Genetic Toxicology, the Key Laboratory of Modern Toxicology of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China;
\textsuperscript{b}Department of Environmental Genomics, Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Cancer Center, Nanjing Medical University, Nanjing, Jiangsu 211166, China;
\textsuperscript{\textcopyright}Department of Epidemiology, the University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA.

Received 20 March 2013, Accepted 27 March 2013, Epub 16 April 2013

Abstract

Although tobacco and alcohol consumption are two common risk factors of head and neck cancer (HNC), other specific etiologic causes, such as viral infection and genetic susceptibility factors, remain to be understood. Human DNA is often damaged by numerous endogenous and exogenous mutagens or carcinogens, and genetic variants in interaction with environmental exposure to these agents may explain interindividual differences in HNC risk. Single nucleotide polymorphisms (SNPs) in genes involved in the DNA damage-repair response are reported to be risk factors for various cancer types, including HNC. Here, we reviewed epidemiological studies that have assessed the associations between HNC risk and SNPs in DNA repair genes involved in base-excision repair, nucleotide-excision repair, mismatch repair, double-strand break repair and direct reversion repair pathways. We found, however, that only a few SNPs in DNA repair genes were found to be associated with significantly increased or decreased risk of HNC, and, in most cases, the effects were moderate, depending upon locus-locus interactions among the risk SNPs in the pathways. We believe that, in the presence of exposure, additional pathway-based analyses of DNA repair genes derived from genome-wide association studies (GWASs) in HNC are needed.

Keywords: DNA repair, molecular epidemiology, head and neck cancer, genetic variation, genome-wide association study

INTRODUCTION

Head and neck cancer (HNC), of which the vast majority are squamous cell carcinomas of the head and neck (SCCHN), is the sixth most common cancer and the seventh leading cause of cancer-related deaths worldwide\textsuperscript{[1]}. In the United States, it has been estimated that 52,610 new cases and 11,500 deaths from HNC may have occurred in 2012, which accounts for 3% of all malignancies\textsuperscript{[2]}. Anatomically, HNC includes cancers of the oral cavity, oropharynx, hypopharynx, and larynx. HNC is considered a complex disease because

\textcopyright 2013 by the Journal of Biomedical Research. All rights reserved.

\begin{thebibliography}{99}
\end{thebibliography}
both genetic and environmental risk factors contribute to its etiology. Epidemiological studies have demonstrated that tobacco use and alcohol consumption are the most known risk factors for HNC\textsuperscript{[5,6,7,8]}. Furthermore, infection with high-risk types of human papilloma viruses (HPVs) has also been recognized as an increasingly important risk factor for HNC, especially for oropharyngeal squamous cell carcinoma\textsuperscript{[5,6,8]}. Family history of cancer is another important risk factor for HNC, suggesting that genetic factors may contribute to HNC susceptibility\textsuperscript{[7]}. DNA is constantly damaged by endogenous oxygen free radicals from metabolism and exogenous (both chemical and physical) mutagens. The DNA damage repair process is a crucial mechanism to maintain the stability of genetic materials in mammalian cells, and different repair pathways are available to repair different types of DNA damage\textsuperscript{[8]}. If this mechanism fails, however, unrepaired damage can result in apoptosis or otherwise may lead to mutation fixation, unregulated cell growth and cancer\textsuperscript{[9]}. There are several DNA repair pathways in human cells, including base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), double-strand break repair (DSBR), and direct reversion repair (DRR)\textsuperscript{[10]}. Depending upon the nature of DNA damaging agents, different repair pathways involving distinct genes could be activated. For example, the BER pathway is one of the important mechanisms responsible for the removal of the oxidative and hydrolytic decay of DNA damage. The damaged base is recognized and cleaved from the DNA backbone to generate an apurinic/apyrimidinic (AP) site, and the abasic site is then restored by apurinic/apyrimidinic endonuclease 1 (APE1); the oxidized base 8-oxoguanine is excised by 8-oxoguanine DNA glycosylase (OGG1), and then the DNA synthesis and ligation are excised by polymerase. These processes require the participation of XRCC1, ADPRT and other proteins, such as DNA ligase III (LIG3)\textsuperscript{[10,11]}. The NER pathway identifies and repairs bulky DNA lesions caused by environmental agents, such as ultraviolet light (UV) radiation and mutagenic chemicals\textsuperscript{[11]}. In general, NER has four steps in the repair activity as follows: (a) damage recognition by a complex of bound proteins including XPC; (b) unwinding of the DNA by the TFIIH complex that includes XPD; (c) removal of the damaged single-stranded fragment by molecules including an ERCC1/XPF complex; and (d) synthesis by DNA polymerases\textsuperscript{[12]}. The DSBR pathway involves two principal pathways, namely homologous recombination (HR) and non-homologous end joining (NHEJ). HR repairs DSBs with a template, such as a sister chromatid or homologous chromosome, whereas NHEJ (including XRCC4, XRCC5, XRCC6, and LIG4) does not require any repair template at all\textsuperscript{[13]}. HR can result in deletions and rearrangements, which involves many molecules, including XRC2, XRCC3, and RAD51. The MMR pathway corrects the errors that occur during DNA replication and recombination and maintains DNA homeostasis. Loss of MMR function results in the accumulation of potential mutations and genetic instability\textsuperscript{[14]}. The major components involved in this pathway include MLH1, MSH2, MSH6, and PMS2. The direct reversal of DNA damage can activate O\textsuperscript{6}-methylguanine methyltransferase (MGMT), which removes a methyl group from the O\textsuperscript{6}-position in guanine and transfers it to its own cysteine residue at codon 145 in the protein\textsuperscript{[15]}. It has been shown that a DNA repair phenotype is genetically determined and thus varies in the general population. For example, our studies have shown interindividual variations of DNA repair capacity (DRC) in the general population and the effect of a suboptimal DRC on the risk of HNC\textsuperscript{[16,17]}. One of the genetic variations is single nucleotide polymorphism (SNP), which may be located in coding regions of genes, non-coding sequences of genes, or in the intergenic regions between genes. SNPs in the coding region can be synonymous, where no amino acid substitution is observed, or non-synonymous, where the base substitution results in an alteration in the amino acid codes that could affect the structure and function of the proteins. For example, SNPs in the 3\textsuperscript{'} and 5\textsuperscript{'} untranslated regions (UTRs) of genes can alter the binding of transcriptional factors, microRNA, and mRNA stability\textsuperscript{[18]}. In the past few years, DNA repair gene SNPs have been extensively studied in the etiology of HNC and found to be associated with HNC risk\textsuperscript{[19]}. In this review, we summarize most of the reported SNPs in DNA repair genes (Table 1), which have been investigated in published epidemiologic studies of HNC.

THE BER PATHWAY

APE1

APE1, also known as APE, APEX, HAP1, and REF-1, plays a central role in the BER pathway by hydrolyzing the phosphodiester backbone immediately at 5\textsuperscript{'}, to the AP site. APE1 can also act as a 3\textsuperscript{'}-phosphodiesterase to initiate repair of DNA single strand breaks that are produced either directly by reactive oxygen species or indirectly through the enzymatic removal of damaged bases\textsuperscript{[20,21]}. A total of 123 SNPs in APE1 have been reported (http://www.ncbi.nlm.nih.gov/projects/SNP/, as of Mar 1, 2013), 18 of which...
are common and potentially functional\cite{22}, but only -656T>G (rs1760944) and Asp148Glu (rs3136820) are the most extensively studied SNPs. Studies on the functional impact of the -656T>G SNP showed that this SNP influenced the transcriptional activity of APE1 and contributed to lung cancer susceptibility\cite{23,24}. As for the Asp148Glu SNP, functional studies suggested that the G variant allele was associated with increased mitotic delay after exposure to ionizing irradiation\cite{25}. Additionally, Li et al.\cite{26} studied the effects

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Reported SNPs</th>
<th>Most studied SNPs</th>
<th>Ref. SNP IDs</th>
<th>Base changes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base excision repair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APE1</td>
<td>123</td>
<td>-656T&gt;G</td>
<td>rs1760944</td>
<td>T&gt;G</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp148Glu</td>
<td>rs3136820</td>
<td>T&gt;G</td>
</tr>
<tr>
<td>ADPRT</td>
<td>1063</td>
<td>Val762Ala</td>
<td>rs1136410</td>
<td>T&gt;C</td>
</tr>
<tr>
<td>OGG1</td>
<td>439</td>
<td>Ser326Gys</td>
<td>rs1052133</td>
<td>G&gt;C</td>
</tr>
<tr>
<td>XRCC1</td>
<td>200</td>
<td>Arg194Trp</td>
<td>rs1799782</td>
<td>C&gt;T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg280His</td>
<td>rs25489</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>LIG3</td>
<td>449</td>
<td>&lt;39G&gt;A</td>
<td>rs2074522</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>PCNA</td>
<td>269</td>
<td>Arg39Ser</td>
<td>rs1050525</td>
<td>A&gt;C</td>
</tr>
<tr>
<td><strong>Nucleotide excision repair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XPA</td>
<td>461</td>
<td>23A&gt;G</td>
<td>rs1800975</td>
<td>A&gt;G</td>
</tr>
<tr>
<td>XPC</td>
<td>692</td>
<td>PAT</td>
<td>-</td>
<td>/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala499Val</td>
<td>rs2228000</td>
<td>C&gt;T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lys939Gln</td>
<td>rs2228001</td>
<td>A&gt;C</td>
</tr>
<tr>
<td>ERCC1</td>
<td>437</td>
<td>Asn118Asn</td>
<td>rs11615</td>
<td>T&gt;C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IVS5+33C&gt;A</td>
<td>rs3212961</td>
<td>C&gt;A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8902C&gt;A</td>
<td>rs3212986</td>
<td>C&gt;A</td>
</tr>
<tr>
<td>ERCC2/XPD</td>
<td>560</td>
<td>Asp312Asn</td>
<td>rs1799793</td>
<td>G&gt;A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lys571Gln</td>
<td>rs1052559</td>
<td>A&gt;C</td>
</tr>
<tr>
<td>ERCC4/XPF</td>
<td>766</td>
<td>Arg415Gln</td>
<td>rs1800067</td>
<td>G&gt;A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser835Ser</td>
<td>rs1798901</td>
<td>C&gt;T</td>
</tr>
<tr>
<td>ERCC5/XPG</td>
<td>742</td>
<td>Asp1104His</td>
<td>rs17655</td>
<td>G&gt;C</td>
</tr>
<tr>
<td><strong>Mismatch repair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLH1</td>
<td>1677</td>
<td>-93G&gt;A</td>
<td>rs1800734</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>MSH2</td>
<td>2635</td>
<td>Gly322Asp</td>
<td>rs487188</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>MSH6</td>
<td>1013</td>
<td>Gly39Glu</td>
<td>rs1042821</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>PMS2</td>
<td>945</td>
<td>Ser775Asn</td>
<td>rs1050960</td>
<td>G&gt;A</td>
</tr>
<tr>
<td><strong>Double-strand break repair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homologous recombination</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XRCC2</td>
<td>661</td>
<td>Arg188His</td>
<td>rs3218536</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>XRCC3</td>
<td>438</td>
<td>Thr241Met</td>
<td>rs861539</td>
<td>C&gt;T</td>
</tr>
<tr>
<td>RAD51</td>
<td>737</td>
<td>155G&gt;C</td>
<td>rs1801320</td>
<td>G&gt;C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>172G&gt;T</td>
<td>rs1801321</td>
<td>G&gt;T</td>
</tr>
<tr>
<td>NBS1</td>
<td>1040</td>
<td>Glu185Gln</td>
<td>rs1805794</td>
<td>G&gt;C</td>
</tr>
<tr>
<td>Non-homologous end-joining</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XRCC4</td>
<td>5014</td>
<td>Ser298Ala</td>
<td>rs3734091</td>
<td>A&gt;C</td>
</tr>
<tr>
<td>XRCC5</td>
<td>1092</td>
<td>69506G&gt;A</td>
<td>rs3835</td>
<td>G&gt;A</td>
</tr>
<tr>
<td>XRCC6</td>
<td>843</td>
<td>-901T&gt;C</td>
<td>rs5751129</td>
<td>T&gt;C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-61C&gt;G</td>
<td>rs2267437</td>
<td>C&gt;G</td>
</tr>
<tr>
<td>LIG4</td>
<td>323</td>
<td>Ile658Val</td>
<td>rs2232641</td>
<td>A&gt;G</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thr9Ile</td>
<td>rs1805388</td>
<td>G&gt;C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp568Asp</td>
<td>rs1805386</td>
<td>T&gt;C</td>
</tr>
<tr>
<td><strong>Direct reversion repair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGMT</td>
<td>5924</td>
<td>Leu84Phe</td>
<td>rs12917</td>
<td>C&gt;T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ile143Val</td>
<td>rs2308321</td>
<td>C&gt;T</td>
</tr>
</tbody>
</table>

of SNPs in APE1, ADPRT, and XRCC1 on SCCHN risk in a hospital-based case-control study and found that no altered risk was associated with the APE1 Asp148Glu SNP in SCCHN. In contrast, Matullo et al. found that the Asp148Glu variant allele was associated with a significantly reduced incidence of upper aero-digestive cancer (UADC), including HNC. A meta-analysis of 28 case-control studies indicated that the Asp148Glu SNP may be a low-penetrance risk factor for cancer development.

ADPRT

ADPRT, also known as PARP1, is involved in DNA damage signaling, genomic stability of damaged cells, BER, recombination, and transcriptional regulation of tumor suppressor genes. Accordingly, activation of ADPRT has been shown in a variety of diseases. It has been reported that 1063 SNPs have been identified in this gene (http://www.ncbi.nlm.nih.gov/projects/SNP/), but a few are common and potentially functional. A Val762Ala SNP (rs11364110) in ADPRT was shown to contribute to significantly lower poly(ADP-ribosyl)ation activities in a dose-dependent manner in response to oxidative damage. In an association study, Li et al. found some evidence of an interaction between the ADPRT 762Ala SNP and variants of other BER genes in the etiology of SCCHN, although this SNP was not found to be associated with risk of other cancers. These discrepant findings across studies may reflect a cancer-specific effect in the etiology of cancers. In addition, previous studies have also indicated that ADPRT expression levels appear to be low in breast cancer, but high in endometrial cancer, which appears to be validated by recently pooled results.

OGG1

The human OGG1 (hOGG1) catalyzes the cleavage of the glycosylic bonds between the modified base and the sugar moiety, leaving an abasic AP site in DNA; the resultant AP site is then incised, and the repair is completed by successive actions of a phosphodiesterase, a DNA polymerase, and a DNA ligase. Among the 439 SNPs identified to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), the most extensively studied Ser326Cys SNP (rs1052133) is located in exon 7 of hOGG1, resulting in an amino acid substitution of serine with cysteine at codon 326. Studies on the functional impact of the hOGG1 Ser326Cys SNP yielded inconsistent results. Many studies reported that no difference was observed in catalytic activities by the Ser326Cys genotype. However, other studies revealed that the Cys allele exhibited reduced DNA repair activity and was associated with risk of cancers. The role of the Ser326Cys SNP and HNC risk was investigated in six studies. The first case-control study, with a small sample size (167 cases and 331 controls), reported a positive association with HNC risk, which was confirmed by Pawlowska et al. Moreover, Sliwinski et al. showed that the Ser326Cys SNP might modify the risk of SCCHN associated with smoking. Nevertheless, Zhang et al. conducted a hospital-based study and found no association between the Ser326Cys SNP and HNC risk, which may be due to the possibility of selection bias of subjects. The discrepant findings in these studies may reflect the differences in genetic background among different ethnic groups under the investigation, carcinogen exposure in different populations, and limited study power as indicated by study sample sizes.

XRCC1

XRCC1, as a scaffold protein for other BER proteins, recognizes single-strand breaks. Although more than 200 SNPs in XRCCI have been identified (http://www.ncbi.nlm.nih.gov/projects/SNP/), only three common SNPs have been widely investigated in cancer risk, which are Arg194Trp (rs1799782), Arg280His (rs25489), and Arg399Gln (rs25487), located in exon 6, 9, and 10, respectively. The Arg399Gln SNP is the most frequently investigated SNP of the BER genes. Several studies reported that the Arg399Gln SNP was associated with risk of various cancers. A meta-analysis conducted by Hu et al. found that the Arg194Trp SNP was associated with a decreased risk of cancer, whereas the Arg280His SNP was associated with an increased risk of cancer. Huang et al. performed a pooled analysis for an association between the Arg399Gln SNP and HNC risk. They found that the XRCC1 399Gln/Gln genotype was associated with a decreased risk among Caucasians, which was supported by other reports that this SNP was associated with measurably reduced DNA repair capacity, as detected by the persistent DNA adducts and elevated levels of sister chromatid exchanges. However, Li et al. failed to identify any association between the Arg399Gln SNP on the risk of SCCHN in a moderately large sample size. They further performed a meta-analysis and found that the Arg399Gln SNP was significantly associated with HNC risk among Asians but not among Caucasians, suggesting that this SNP has an effect or a higher penetrance in Asian populations. As for the Arg194Trp SNP, many studies reported detectable effects on DNA-adduct levels, frequency of mutations, and sensitiv-
SNPs, namely Asn118Asn. However, the function of the Arg280His SNP is still not fully understood. This SNP could produce an amino acid change and decrease genomic stability. A recent meta-analysis suggested that the Arg280His SNP may play a role in cancer development, though only one study on HNC was included in the analysis. Inconsistent with the previous findings, Applebaum et al. reported that the XRCC1 polymorphisms might confer susceptibility to SCCHN in the context of smoking, among those who are seronegative for HPV16.

LIG3 and PCNA

LIG3 encodes ligase III, which is DNA- and ATP-dependent, an essential component in the BER pathway in the repair of damages caused by free radicals and, the levels of damage was found to be also directly associated with the expression levels of XRCC1. Totally, 449 SNPs in LIG3 and 269 SNPs in PCNA have been reported. Although many SNPs in LIG3 and PCNA genes were reported to have been studied in many cancers, such as rs3136025 G>A and rs2074522 G>A in LIG3, and rs1050525 in PCNA, no significant associations were found. Therefore, the associations of LIG3 and PCNA SNPs with HNC risk are still not fully understood.

THE NER PATHWAY

XPA

The XPA gene encodes a zinc-finger DNA-binding protein that participates in DNA excision repair to maintain genomic integrity. The XPA protein is also involved in both global genome and transcription-coupled repair pathways. Of 461 SNPs reported to date, a common 23A>G SNP (rs1800975) in the 5′-UTR has been widely studied, although the functional significance of this SNP is still unknown. It has been demonstrated that the 5′-UTR may regulate gene expression through transcriptional and post-transcriptional control mechanisms, that this SNP may affect the binding between translational factors and the promoter, and that individuals with the G variant had a reduced risk of lung cancer in Caucasian and Korean populations. Previously, five case-control studies with small sample sizes had reported the association between the 23A>G SNP and HNC risk, and the results were inconclusive, partially because of the possible weak effect of the SNP on HNC and the relatively small size in each of the published studies. For example, Sugimura et al. conducted a case-control study with 122 oral squamous cell carcinoma (OSCC) cases and 241 controls and found that the 23A>G SNP was associated with OSCC risk. Another study found similar results, but also the synergistic effects of XPA and XPD SNPs, as well as smoking status, on oral cancer risk. In contrast, Jelonek et al. did not find any association between this SNP and risk of HNC. Later, a meta-analysis on all eligible case-control studies between the 23A>G SNP and HNC risk demonstrated that the G allele was a low-penetrance risk factor for cancer development.

XPC

The XPC gene encodes part of the XPC-HR23B complex that plays a role in the early step of NER by initially detecting the DNA damage. The complex can bind to damaged DNA and change the DNA conformation around the lesion. Among all identified 692 SNPs of XPC, three common biallelic poly (AT) insertion/deletion (PAT, +/−), Ala499Val (rs2228000), and Lys939Gln (rs2228001) have been extensively studied. The XPC-PAT polymorphism, located in intron 9, was found to be in linkage disequilibrium with the Ala499Val SNP, which has been reported to be associated with a decreased DNA repair capacity as measured by chromatid aberrations. However, the functionality of the Lys939Gln SNP is still unknown. Shen et al. firstly reported that the XPC-PAT + allele was associated with the risk of SCCHN in a small study (287 cases and 311 controls). However, in their subsequent large case-control study, they found that only the XPC Ala499Val SNP was associated with a significantly increased risk of SCCHN. Although Yang et al. found that XPC expression may influence SCCHN risk, no effect of XPC-PAT genotypes on SCCHN risk was observed in a Korean population. These findings should be validated in larger studies with the same ethnic background.

ERCC1

ERCC1 is essential for the NER pathway and plays a critical role in genome stability. SNPs in ERCC1 have been reported to contribute to susceptibility to carcinogenesis. Among the 437 SNPs reported, at least three common ERCC1 SNPs, namely Asn118Asn (rs11615), IVS5 + 33C>A (rs3212961), and 8092>A (rs321986), are probably functional. Two meta-analyses revealed that Asn118Asn and IVS5 +33C>A SNPs, but not 8092>C, A, are low-penetrance risk factors for cancer risk. As for HNC, most studies focused on 8092>C SNP and HNC risk. In a small study with 67 SCCHN cases and 73 controls,
Yang et al. studied the association between mRNA expression and the 8092C>A SNP of ERCC1, and found lower ERCC1 mRNA expression in SCCN patients than in controls, but no association of the 8092C>A SNP with the risk of SCCN or expression of ERCC1 mRNA[89]. In a larger study, Sturgis et al. found that the 8092C>A SNP combined with the XPD Asp312Asn SNP may contribute to the risk of SCCN[90]. However, the results have not been replicated by An et al. in a much larger hospital-based case-control study.[83]. In the stratification of tumor site in the meta-analysis by Zhang et al., no significant risk associated with the 8092C>A SNP was found for SCCN in different genetic models[88].

ERCC2/XPD

XPD is an enzyme in the NER pathway that removes certain DNA cross-links, UV photo-lesions, and bulky chemical adducts[91]. Of 560 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), two XPD polymorphic loci that are of most interest in molecular epidemiology studies are Asp312Asn (rs1799793) and Lys751Gln (rs13181). It has been suggested that the XPD 312 Asn variant allele may be associated with reduced repair of aromatic DNA adducts[92]. In Lys751Gln SNP, glutamine is changed to lysine at codon 751, which has been reported to have a reduced DNA repair proficiency[93]. Several studies have investigated the associations of XPD Asp312Asn and Lys751Gln with HNC risk but generated inconsistent results[9,76,77,89,94-97]. Sturgis et al. firstly studied the association between the Asp312Asn SNP and HNC risk and found that this SNP had a significant association with HNC risk when combined with the ERCC1 8092C>A SNP[89]. Other studies also showed that HNC risk was not significantly associated with the Asp312Asn SNP[9,76,77,95,96]. A recent comprehensive meta-analysis of nine case-control studies on XPD Asp312Asn SNP and HNC risk reported that the Asp312Asn SNP might not be associated with HNC risk[86]. As for the Lys751Gln SNP, Yuan et al. performed a pooled analysis of 15 case-control studies to assess the overall association between this SNP and HNC risk and found that a significantly elevated HNC risk was associated with Lys751Gln SNP[99], and this increased risk was more pronounced among Europeans but not among Asians.

ERCC4/XPF

ERCC4 is a key enzyme responsible for excising bulky adducts from damaged DNA because of its unique function in damage site recognition. ERCC4 contains the catalytic domain of the nuclease, and ERCC1 is required for DNA binding and stabilization of ERCC4[88]. To date, 766 SNPs in ERCC4 have been reported (http://www.ncbi.nlm.nih.gov/projects/SNP/) and subsequent association studies found that some of them are likely to contribute to susceptibility to several kinds of cancers, including HNC, but only two common SNPs Arg415Gln (rs1800067) and Ser835Ser (rs1799801) in the exons of ERCC4 have been extensively studied in the epidemiologic investigations[76,93,100]. Although the results from these three reported studies are not consistent, a recent meta-analysis indicated that these two SNPs were not associated with HNC risk. Some publication bias was observed in the meta-analysis, because the results were not adjusted for other covariants, including age, smoking status, drinking status, and other environmental factors[101].

ERCC5/XPG

ERCC5 has structure-specific endonuclease activity that is essential for two incision steps in NER[102]. Of 742 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), the Asp1104 His SNP (rs17655), located in exon 15 of XPG, has been largely studied for susceptibility to HNC. Although the His1104Asp SNP together with SNPs of several other NER genes may jointly contribute to the variability of the DRC phenotype, all published data did not show any significant association with risk of HNC[76,103-105].

THE MMR PATHWAY

MLH1 and MSH2

MLH1 and MSH2 are the key components of the MMR system that participates in the recognition of nucleotide mismatches occurring during DNA replication and in the recruitment of additional repair proteins to correct replication errors. Numerous SNPs have been identified in MLH1 and MSH2 (1677 and 2635, respectively, http://www.ncbi.nlm.nih.gov/projects/SNP/), but few of which have been investigated for their associations with cancer. For example, the MLH1 -93G>A SNP (rs1800734) is located in the promoter region that is responsible for maximal transcriptional activity of this gene[106,107]. The association between the -93G>A SNP and cancer risk was investigated most in colorectal cancer[108,109]. Reduced expression levels of MSH2 have been reported in hereditary nonpolyposis colon cancer (HNPCC) and some other human cancers[106-113]. The Gly322Asp SNP is located in the coding region of the MSH2 gene, resulting in a modest decrease in mismatch repair efficiency[114,115]. However, no study on the association of both the -93G>A and Gly322Asp SNPs with HNC risk has been reported.
MSH6 and PMS2

The MMR proteins MSH6 and PMS2 have been shown to interact with MSH2 and MLH1, respectively\(^{[116]}\), and the alteration in the expression of MSH6 can perturb MMR, while the protein encoded by PMS2 plays an essential role in repairing DNA by forming an active protein complex with MLH1 that interacts with MSH2 bound to mismatched bases. Although many SNPs have been identified in these genes (1013 for MSH6 and 945 for PMS2, http://www.ncbi.nlm.nih.gov/projects/SNP/), inherited mutations are the most studied topics for colorectal cancer. The mutations of PMS2 were firstly found in a HNPCC patient\(^{[117]}\), and inherited mutations in MSH2 and MLH1 are the most prevalent cause of colorectal tumors or HNPCC, as these tumors typically have microsatellites, a marker for malfunction of MMR\(^{[118]}\), which were not common in HNC.

THE DSBR PATHWAY

Homologous recombination (HR)

XRCC2

The XRCC2 gene, located at 7q36.1, is an essential part of the HR repair pathway and also a functional candidate for its involvement in cancer progression, including HNC. Among the 661 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), a common SNP Arg188His (rs3218536) in XRCC2 has been reported in epidemiological studies. Benhamou et al. found a significant association between the Arg188His SNP and risk of upper aerodigestive tract (UADT) cancers\(^{[119]}\); in particular, the 188His allele was associated with a significantly increased risk of pharyngeal cancer. Furthermore, this SNP was reported to be associated with an increased risk of oral cancer risk\(^{[120]}\). In contrast, the Arg188His and three other SNPs in the promoter region have been associated with reduced risk in other cancers, such as bladder cancer\(^{[121]}\).

XRCC3

XRCC3, involved in the HR of DNA DSBR and cross-links, is a member of an emerging family of RAD51-related proteins that likely participate in HR to maintain genomic stability and repair DNA damage\(^{[122]}\). XRCC3 deficient cell lines display reduced HR. Of the 438 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), the main Thr241Met (rs861539) SNP in XRCC3 leads to an amino acid substitution at codon 241, which has been reported to be associated with slightly but not significantly decreased DNA repair capacity\(^{[123]}\). Recently, a meta-analysis of 15 eligible studies with 3,191 cases and 5,090 controls found some evidence for an association between the Thr241Met SNP and HNC risk\(^{[124]}\). However, another meta-analysis by Flores-Obando et al. revealed that the SNP did not contribute to the risk of HNC\(^{[125]}\). These two different results may be due to the number of studies included in the analyses (15 vs. 10) with various study powers.

RAD51

RAD51 is involved in HR and DSBR in DNA and DNA cross-links and for the maintenance of chromosome stability\(^{[126]}\). RAD51 is associated with BRCA1 and BRCA2 tumor suppressor gene products, suggesting that a defect in recombination may lead to tumor development\(^{[127]}\). Of the 737 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), two SNPs in the 5’-UTR of RAD51, 135G>C (rs1801320) and 172G>T rs1801321, have been investigated for their associations with the risk of sporadic breast cancer or breast cancer in non-BRCA1/2 carriers, though the results are yet controversial and inconclusive\(^{[128]}\). There are only few studies investigating the association of the 135G>C SNP as a risk factor for HNC so far, but they reported inconsistent findings of either increased, decreased or lack of any HNC risk associated with the 135C allele\(^{[129,130]}\). A small study of 81 HNC cases and 111 controls also suggested that no significant effect of this SNP and HNC risk was observed, which was modulated by smoking status and gene-gene interactions\(^{[131]}\). Moreover, Lu et al. observed that the 172G>T SNP was associated with significantly reduced risk of SCCHN\(^{[132]}\), which was confirmed by Gresner et al.\(^{[133]}\). A possible mechanism underlying the protective effect of the 172G>T SNP in HNC may be due to increased cellular apoptotic potential, or the capacity of the DNA DSB repair system.

NBS1

NBS1 is part of a protein complex that forms in response to DNA damage to maintain chromosomal integrity. The exact role of NBS1 in DNA repair is not fully understood, because NBS1 does not have a DNA binding site or kinase activity. As one of the most commonly studied SNPs, among the 1040 reported SNPs (http://www.ncbi.nlm.nih.gov/projects/SNP/), the Glu185Gln SNP is non-synonymous, with an amino acid change that may affect the function of NBS1 and the protein-protein interaction\(^{[133]}\). Zheng et al. observed a significant difference in genotype frequencies of the Glu185Gln SNP between nasopha-
ryngal carcinoma cases and controls and found that the C allele was associated with the risk of invasive or metastatic disease, compared with the G allele\textsuperscript{134}.  

**Non-homologous end-joining (NHEJ)**

**XRCC4**

XRCC4, a key component of the NHEJ repair pathway, is found to restore DNA DSB repair. It is basically required for precise end-joining of blunt DNA DSBs. There are 5014 SNPs that have been identified to date (http://www.ncbi.nlm.nih.gov/projects/SNP/). Of these SNPs, Tseng et al. reported that the Ser298Ala (rs3734091) SNP in exon 3 of XRCC4 was associated with the risk of oral cancer in Taiwanese populations\textsuperscript{135} and found that individuals with 247A allele had a 2.04-fold higher risk of cancer. In addition, Chiu et al. firstly observed that an intron 3 del genotype (rs28360071) may be associated with oral cancer in Taiwanese patients\textsuperscript{136}.  

**XRCC5 and XRCC6**

XRCC5 (also named Ku80 or Ku86) and XRCC6 (also known Ku70), are the components of the major repair pathway of DSBs in eukaryotic cells during most phases of the cell cycle, particularly the G0/G1 phases\textsuperscript{137}. These proteins are essential for the maintenance of chromosomal integrity and the viability of cell and organism, as NHEJ plays a more predominant role in adult mammalians than the alternative DSB repair mechanism, homologous recombination\textsuperscript{138,139}. There are 1092 and 843 SNPs that have been identified to date for XRCC5 and XRCC6, respectively (http://www.ncbi.nlm.nih.gov/projects/SNP/). A small study (152 SCCHN patients and 157 controls) found that the genotypes of rs3835 in XRCC5 and rs2267437 in XRCC6 were neither associated with SCCHN risk, nor was gene-smoking or gene-alcohol interaction observed\textsuperscript{136}. Bau et al. found that the XRCC6 promoter SNP rs5751129 was associated with oral cancer risk, but not rs2267437, rs132770, or rs132774\textsuperscript{140}. The same group reported that the XRCC5 rs828907, but not rs11685387 or rs9288518, was associated with oral cancer susceptibility\textsuperscript{141}. In addition, Yang et al. investigated subjects with oral premalignant lesions in a case-control study and found no association between the XRCC5 rs1051685 SNP and the risk of oral cancer\textsuperscript{142}.  

**LIG4**

LIG4 jointed the final rejoining step of NHEJ with XRCC4, and its deficiency leads to significant genomic instability\textsuperscript{143}. Among the 323 SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), an Ile658Val (rs2232641) SNP in the exon of LIG4, affecting LIG4 function, has been found to be slightly associated with the developing of cervical carcinoma in a Chinese population\textsuperscript{144}. However, no published studies have investigated the association between the Ile658Val SNP and HNC risk. Werbrouck et al. investigated the role of another two SNPs, Thr9Ile (rs1805388) and Asp568Asp (rs1805386), in SCCHN development and found that the Thr9Ile SNP was associated with a significantly reduced risk for SCCHN\textsuperscript{140}.  

**THE DR PATHWAY**

**MGMT**

The only known enzyme in the DR pathway is MGMT that removes a methyl group from the O6-position in guanine and transfers it to its own cysteine residue at codon 145 in the protein, inactivating the MGMT protein itself while repairing guanine. It is reported that the MGMT knockout mice had a higher incidence of nitrosamine-induced tumorigenesis\textsuperscript{145}. Of the SNPs reported to date (http://www.ncbi.nlm.nih.gov/projects/SNP/), two functional common SNPs, Leu84Phe (rs12917) and Ile143Val (rs2308321), have been extensively investigated for their roles in HNC. Recently, Zhang et al. extended the SNPs in MGMT including the SNPs in the promoter region and found that none of SNPs in MGMT had a substantial effect on SCCHN risk, but a joint effect of several MGMT variants may contribute to risk and progression of SCCHN\textsuperscript{146}. Furthermore, they did a mini-pooled analysis of previously published studies on the Leu84Phe SNP and HNC risk\textsuperscript{146} and found that this single variant was not associated with risk of SCCHN.  

**GWAS IN HNC**

During the past few years, GWASs have identified a large number of robust associations between specific chromosomal loci and complex human diseases\textsuperscript{147}. For HNC, a recent GWAS identified two novel variants (rs1494961 at 4q21 and rs4767364 at 12q24) and three remaining variants (rs1573496, rs1229984, and rs698) located in the ADH gene cluster that is associated with risk of UADT cancers in Caucasians\textsuperscript{148}. Yuan et al. performed a validation study of GWAS findings for their associations with HNC risk in a hospital-based case-control study and found that rs1229984 at 4q23 was associated with significantly increased risk of HNC in a Chinese population\textsuperscript{150}. However, the biological relevancy of
GWAS discoveries is still largely unknown. Many of the disease associations identified by GWAS do not involve previously known candidate genes, such as DNA repair pathway genes, and many associated markers are in genomic regions harboring unknown genes. Therefore, the functions of those variants identified in GWAS should be validated via biological assays, pathway analysis or bioinformatics in the post-GWAS era.

**CONCLUSION**

DNA repair mechanisms play an important role in correcting DNA damage and preserving genomic integrity, and it is obvious that abnormalities of these mechanisms are strongly associated with cancer risk, including HNC. Specifically, genetic variants in DNA repair genes (Table 1) are associated with alteration of an individual’s susceptibility to malignancy. Here, we have summarized most of previously published studies that have investigated associations between SNPs in DNA repair genes and HNC risk. It is interesting that the only GWAS performed in Caucasian populations did not identify any DNA repair gene SNPs as the top-hits. More GWAS of HNC in different ethnic groups are needed to provide the whole picture of the importance of the SNPs in these genes. Many association studies have focused on the roles of those common and functional SNPs in the BER, NER, MMR, DSBR, and DR pathways. Among the above-reviewed genes, little is known about the prognostic and therapeutic values in HNC. Therefore, further investigations of alterations in DNA repair pathways are needed, which are essential to identify their clinical application in HNC.

**References**


[22] Xi T, Jones IM, Mohrenweiser HW. Many amino acid
substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. *Genomics* 2004; 83: 970-9.


DNA repair gene polymorphisms and HNC


[75] Bau DT, Tsai MH, Huang CY, Lee CC, Tseng HC, Lo


