Published in final edited form as:

Cancer Epidemiol. 2011 April ; 35(2): 175–181. doi:10.1016/j.canep.2010.08.007.

Association studies of *Excision Repair Cross-Complementation Group 1 (ERCC1)* haplotypes with lung and head and neck cancer risk in a Caucasian population

Nathan R. Jones^{a,c}, Thomas E. Spratt^{b,d}, Arthur S. Berg^{a,e}, Joshua E. Muscat^{a,e}, Philip Lazarus^{a,c,e}, and Carla J. Gallagher^{a,e}

^aMolecular Epidemiology and Cancer Control, Penn State Cancer Institute, Penn State College of Medicine, Hershey, PA

^bChemical Carcinogenesis and Chemoprevention Programs, Penn State Cancer Institute, Penn State College of Medicine, Hershey, PA

^cDepartment of Pharmacology, Penn State College of Medicine, Hershey, PA

^dDepartment of Biochemistry and Molecular Biology, Penn State College of Medicine, Hershey, PA

^eDepartment of Public Health Sciences, Penn State College of Medicine, Hershey, PA

Abstract

Background—The formation of bulky DNA adducts caused by diol epoxide derivatives of polycyclic aromatic hydrocarbons has been associated with tobacco-induced cancers, and inefficient repair of such adducts by the nucleotide excision repair (NER) system has been linked to increased risk of tobacco-induced lung and head and neck (H&N) cancers. The human excision repair cross-complementation group 1 (ERCC1) protein is essential for a functional NER system and genetic variation in *ERCC1* may contribute to impaired DNA repair capacity and increased lung and H&N cancer risk.

Methods—In order to comprehensively capture common genetic variation in the *ERCC1* gene, Caucasian data from the International HapMap project was used to assess linkage disequilibrium and choose four tagSNPs (rs1319052, rs3212955, rs3212948, and rs735482) in the *ERCC1* gene to genotype 452 lung cancer cases, 175 H&N cancer cases, and 790 healthy controls. Haplotypes were estimated using expectation maximization (EM) algorithm, and haplotype association with cancer was investigated using Haplo.stats software adjusting for known covariates.

Results—The genotype and haplotype frequencies matched previous estimates from Caucasians. There was no significant difference in the prevalence of rs1319052, rs3212955, rs3212948, and rs735482 when comparing lung or H&N cancer cases with controls (p-values > 0.05). Similarly, there was no association between *ERCC1* haplotypes and lung or H&N cancer susceptibility in

© 2010 Published by Elsevier Ltd. All rights reserved.

Please address correspondence to: Carla Gallagher, Ph.D., Department of Public Health Sciences, Penn State College of Medicine, 500 University Drive, CH69, Hershey, PA; fax: (717) 531-0480; cgallagher@hmc.psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement

The authors of this work have no conflicts of interest to disclose.

this Caucasian population (p-values > 0.05). No associations were found when stratifying lung cancer cases by histology, sex, smoking status, or smoking intensity.

Conclusions—This study suggests that *ERCC1* polymorphisms and haplotypes do not play a role in lung and H&N cancer susceptibility in Caucasians.

Keywords

ERCC1; haplotypes; lung cancer; head and neck cancer; Caucasian; smoking; genotyping

Introduction

The formation of bulky DNA adducts caused by various compounds such as the diol epoxide derivatives of polycyclic aromatic hydrocarbons (PAHs) are likely causes of tobacco-induced cancers [1]. These DNA adducts are mutagenic and are repaired by the nucleotide excision repair (NER) system. A sub-optimal DNA repair capacity has been shown to be a risk factor for tobacco-induced cancers [2,3]. Excision repair cross-complementation group 1 protein (ERCC1) forms a heterodimer with XPF that cleaves the DNA 5' to the lesion [4] and provides a link to the NER machinery through its interaction with the XPA protein [5] as well as binding to the DNA [6]. XPF-ERCC1 has other DNA repair functions; it is involved in homologous recombination [7] as well as the repair of interstrand crosslinks [8,9]. Since ERCC1 is essential for a functional NER system [4], alterations in the expression or activity of this protein may affect NER-related DNA repair capacity.

Polymorphisms in DNA repair genes have been associated with increases in cancer susceptibility. Numerous case-control studies have investigated *ERCC1* polymorphisms and their potential association with several types of cancer [10–17]. The 3'UTR C8092A and synonymous exon 4 T19007C polymorphisms have been studied most extensively since there are not any non-synonymous SNPs in the coding region. The 3'UTR C8092A polymorphism is thought to be associated with alterations in transcript stability [11,17] while the T19007C polymorphism has been associated with altered mRNA levels [11]. Low expression of *ERCC1* mRNA has been associated with increased lung and H&N cancer susceptibility [15,18], while high expression has been shown to improve DNA repair capacity [19]. The evidence that *ERCC1* mRNA levels have an effect on NER capacity has been further demonstrated by studies linking high *ERCC1* expression in tumor tissue with resistance to DNA cross-linking agents such as cisplatin [20–24].

There is conflicting evidence for the association of *ERCC1* SNPs with lung cancer risk [11,12,25,26], and a meta-analysis of the C8092A and T19007C polymorphisms found no association [27]. There are fewer studies involving *ERCC1* polymorphisms and H&N cancer risk, but two studies suggest that the C8092A polymorphism increases risk [28,29]. Because of its potential as an important genetic biomarker for cancer risk, other approaches such as haplotyping are necessary to examine the role of *ERCC1* genetics on tobacco-related cancer risk. Haplotype analysis allows for the comprehensive analysis of genetic variations across an entire gene as it captures common genetic variation within a given haplotype block without requiring prior knowledge of SNP functionality. Despite the overall lack of association with known functional SNPs, Ma et al. found a protective effect of the TCCCATT haplotype and lung cancer risk in a large Chinese population [30]. To our knowledge, there are currently no association studies involving *ERCC1* haplotypes and H&N cancer risk and no studies have been performed examining *ERCC1* haplotypes and cancer risk in a Caucasian population. In the present study, we investigated four tagSNPs

that represent the common genetic variation in the *ERCC1* promoter and coding region to test for potential associations with lung and/or H&N cancer risk in Caucasians.

Materials and Methods

Study Population

The case-control study was conducted at the H. Lee Moffitt Cancer Center (Tampa, FL) from 2000 to 2003. Caucasian lung and H&N cancer cases ($n = 452$ and $n = 175$, respectively) were newly-diagnosed subjects with histologically confirmed lung or H&N cancer with no past history of other tobacco-related cancers. Caucasian controls ($n = 790$) were selected from community residents attending the Lifetime Cancer Screening facility of the Moffitt Cancer Center. Control subjects were randomly selected from thousands of community residents who underwent prostate-specific antigen testing, skin examinations, endoscopy, or mammography. Spiral computed tomography for lung cancer was not done at the clinic. The Lifetime Cancer Screening facility conducts community outreach and education programs throughout the Tampa Bay area, including lecture series, screening events, health fairs, literacy programs, and community-based partnerships. A list of control IDs was matched against the hospital patient database to identify any subject who might have developed cancer. All control subjects with a new cancer diagnosis were excluded from this study. Ninety-nine percent of the hospital cancer patients and ninety-seven percent of the clinic control patients who were asked to participate in the study consented. All study subjects signed a consent form approved by the institutional review board. A trained interviewer administered a structured questionnaire that obtained lifestyle and smoking history information including levels of education, occupation, year of smoking onset, current smoking status, number of cigarettes smoked per day, and years since quitting (for former smokers). The questionnaire was an abbreviated form of a previously-validated questionnaire used by investigators at the American Health Foundation in a large hospital-based, case-control study involving over 10,000 cases and 30,000 controls [31]. Individuals were categorized as current, former, or never smokers for the discrete variable 'current smoking status' according to the following method of classification: current smokers had smoked at least one or more cigarettes per day for the past year, former smokers had quit smoking one or more years prior to the interview, and never smokers had smoked fewer than 100 cigarettes in their lifetime. The medical charts of the case subjects were reviewed to obtain diagnostic and pathology records.

SNP selection & genotyping

CEPH genotypes representing individuals with European ancestry were downloaded from the International HapMap Project [32] and linkage disequilibrium (LD) was determined using Haploview software [33]. LD was estimated between all pairs of SNPs using the D' statistic. Haplotype block structure was determined using the Solid Spine of LD option of Haploview with the block extended if pairwise D' between SNPs was greater than 0.80.

According to HapMap data, the coding region and promoter of the *ERCC1* gene lie in separate haplotype blocks (Figure 1). HapMap data was used to identify four haplotypes in the coding region of the *ERCC1* gene, which can be distinguished by three tagSNPs (rs3212948, rs3212955, rs735482), and two haplotypes in 10 kb of the promoter region, which can be distinguished by one tagSNP (rs1319052). All four tagSNPs had >15% minor allelic frequency (MAF). TagSNPs were genotyped using pre-designed Taqman 5'-exonuclease genotyping assays, according to manufacturer's instructions and SDS 2.2.2 software was used for automated determination of genotypes (Applied Biosystems, Foster City, CA). Briefly, allele-specific probes for rs1319052, rs3212955, rs3212948, and rs735482 were labeled with the fluorescent dyes VIC and FAM. During extension, the 5'-

exonuclease activity of the Taq polymerase separates the fluorophore from the non-fluorescent quencher. A post-amplification allelic discrimination run on the ABI 7900HT was used to determine genotype based on the relative amount of fluorescence of VIC and FAM. PCR reactions were carried out in a total reaction volume of 5 μ l in 384-well plates using the ABI 7900 HT Sequence Detection System, and the thermal cycling conditions were 50°C for 2 min, 95 °C for 10 min, and then 40 cycles at 95 °C for 15 s and at 60 °C for 90s. Individuals involved in genotyping were blind to patient status.

Statistical Analysis

Hardy-Weinberg equilibrium was assessed by χ^2 tests. Odds ratios (OR) and 95% confidence intervals (CI) of association between cancer and SNP genotype were estimated using logistic regression. Potential confounding of the association between genotype and cancer risk by known risk factors was explored using Spearman rank correlation analyses and multivariate logistic regression models, including stepwise regression models. Single SNP associations were analyzed using SPSS 15.0 (SPSS Inc., Chicago IL).

Haplotypes were estimated using the expectation maximization (EM) algorithm and haplotype association with cancer was investigated using Haplo.stats software [34], adjusting for known covariates (age, gender, and pack-years). In addition to being analyzed in its entirety, the lung cancer data set was stratified by histology, gender, and smoking status to search for associations. Stratified analyses were not performed in the H&N cancer data set due to limited power. Significance of associations was calculated by a 1,000-permutation algorithm. Rare haplotypes (<2%) were excluded from the analysis.

The power calculation to identify a genotypic effect in the case control studies was performed with the software PS version 3.0.14 [35] and was determined for a range of genotypic frequencies (0.1, 0.2 and 0.4) and odds ratios (1.6, 1.9, and 2.2) using a case-control design. The power calculations were based on two-sided tests with a 0.05 false positive level. With the given sample sizes, there was sufficient power to detect an increased risk in the lung and H&N cancer studies over nearly the entire range of frequencies and effect sizes. Specifically, to detect an OR $\geq$ 1.9 in the H&N study, there was 92% power for a SNP with a MAF of 0.2 and 97% power for a SNP with a MAF of 0.4. The power was even greater in the lung cancer study due to larger sample sizes. The SNPs observed in this study had high MAFs in the controls suggesting there was sufficient power to observe a significant effect if the SNP contributed an increased risk with an odds ratio of at least 1.6.

Results

The basic demographic profile of the lung cancer case control subjects is shown in Table 1. Fifty-four percent were men and 46% were women. The mean ages of cases and controls were 64.2 ± 9.8 and 58.4 ± 10.4 y, respectively ($p < 0.01$). A significantly ($p < 0.01$) higher percentage of lung cancer cases than controls ever smoked cigarettes (92% vs. 62%). The mean pack-years of smoking for lung cancer cases (56 pack-years) was significantly ($p < 0.01$) higher than the mean pack-years observed for controls (24 pack-years). A significantly ($p < 0.01$) lower percentage of lung cancer cases (21%) were college educated as compared to controls (39%), and the mean BMI of 26.9 ± 5.1 in lung cancer cases was similar to that observed in controls (27.2 ± 4.9). Among the lung cancer case group, the most frequent tumor histology was adenocarcinoma (38%), followed by squamous cell carcinoma (24%).

The demographic profile of the H&N cancer cases (Table 2) shows that a significantly higher percentage of men ($p < 0.01$) were found in the cases (74%) versus controls (54%). The mean age of 58.5 ± 11.4 y for H&N cancer cases was similar to that observed for controls. A significantly ($p < 0.01$) higher percentage of H&N cancer cases than controls ever

smoked cigarettes (75% versus 62%) and the mean pack-years for H&N cancer cases (48 pack-years) was significantly higher ($p < 0.01$) than that observed for controls (24 pack-years). A significantly ($p < 0.01$) lower percentage of H&N cancer cases (22%) were college educated and the mean BMI of 26.7 ± 5.5 in H&N cancer cases was similar to that observed in controls.

All SNPs were found to be consistent with Hardy-Weinberg equilibrium. Of 376 samples repeated for quality control purposes, two samples (~0.5%) yielded contradictory results. The MAF of the tagSNP genotypes in controls were 0.40 (rs1319052), 0.39 (rs3212948), 0.25 (rs3212955), and 0.15 (rs735482), which were similar to that observed in the Caucasian data from HapMap.

No association was found between individual SNPs and lung or H&N cancer risk (Table 3 and Table 4). The genotype frequencies, OR, and 95% CI are shown, and indicate that none of the SNPs approached significance. Odds ratios were calculated by logistic regression adjusting for the covariates age, gender, and pack-years. When comparing the H&N and lung cancer cases, it is apparent that no trend exists for any of the SNPs. The genotypes that appear to be modestly protective in one set of cases are associated with risk in the other set of cases, and vice versa. For example, in the lung cancer cases the OR for the rs1319052 G/A genotype is 0.82 (95% CI = 0.61–1.11); however, in the H&N cases the OR for this genotype is 1.29 (95% CI = 0.88–1.89). We used three of the tagSNPs to distinguish four haplotypes in the *ERCC1* coding region and one tagSNP to distinguish two haplotypes in the promoter region of the *ERCC1* gene. In this study, no association was found between any of the *ERCC1* haplotypes and lung or H&N cancer risk, adjusting for age, gender, and pack-years (Table 5 and Table 6). Lung cancer cases were stratified by histology, gender, smoking status, and smoking intensity in an attempt to identify any specific associations that might be present, but none were found. In addition, non-small cell lung cancer and small cell lung cancer cases were analyzed separately because the etiology of these diseases may be different regarding *ERCC1* status, but no associations were found. Based on the high linkage disequilibrium between all four SNPs in our population, an alternate haplotype combination including all four SNPs was analyzed, and no association with lung or H&N cancer risk was observed (data not shown).

Discussion

Understanding the genetics of lung cancer etiology will enable individuals with high-risk genotypes to be targeted for screening and intervention. Although smoking cessation is currently the only intervention that has been shown to be effective in reducing risk of lung cancer in humans, several chemoprevention agents have shown promise in murine models [36–38]. While it remains to be seen whether chemoprevention of lung cancer can be achieved in humans, genetic screening can help physicians assess an individual's risk more accurately and target them for early detection and lung cancer screening [39].

The results of this study suggest that *ERCC1* polymorphisms do not play a role in lung or H&N cancer susceptibility in Caucasians. Previous studies investigating the association of *ERCC1* polymorphisms with tobacco-related cancers have shown mixed results. For example, Zienoldiny et al found a protective effect for the T19007C (C/C) genotype on lung cancer risk in a Norwegian population [11]. However, studies of a large American population, a Danish population, and a Chinese population reported no association with lung cancer risk for this SNP [12,25,26]. Sugimura et al found a significant gene-smoking interaction between the C8092A polymorphism and oral cancer risk in a Japanese population [29]. However, no significant association was observed between the C8092A (C/C) genotype and risk of squamous cell carcinoma of the head and neck in a small Caucasian

population, though a modestly significant increase in risk ($p=0.04$) was seen in individuals that also had the *ERCC2* G23591A (A/A) genotype [28].

More recently, Ma et al. found a protective effect for the C allele of rs3212948 in a Chinese population, in which the C allele is the minor allele (MAF=22%) [30]. In the current study, this association was not observed for the rs3212948 C allele and lung or H&N cancer risk. Instead, a trend towards increased risk was noted for the C allele in both cancer populations. Since it is the major allele in Caucasians, sample size was not a likely contributing factor to this contradictory finding. Given the broad genomic differences between these races, it is not uncommon to find differences in association between them. In addition, rs3212948 is an intronic SNP and is not expected to be the functional SNP. Instead, the association is expected to be due to high LD with the functional SNP, which was not identified in the Chinese study [30]. It is possible that Caucasians lack the functional SNP, though the protective effect was only seen in heterozygotes, which is a difficult phenomenon to explain [30]. In the current study, there was no effect of the heterozygous (C/G) genotype for rs3212948 on lung or H&N cancer risk in Caucasians. The rs1319052, rs3212955, and rs735482 SNPs also were not associated with lung or H&N cancer risk in our population.

The *ERCC1* gene is located in the chromosomal region 19q13.3, and haplotypes of chromosome 19q13.2–3 have been associated with cancer risk in previous studies [25,40–43]. With respect to tobacco-related cancers, a haplotype of polymorphisms encompassing the *ASE-1*, *RAI*, and *ERCC1* genes was associated with lung cancer risk in a large Danish cohort study [25,42]. The association was strongest in middle-aged women that smoked heavily, and the authors suggested that a functional polymorphism in the *RAI* gene was most likely to be responsible for the observed association. Yin et al found no association between the silent *ERCC1* Asn118Asn SNP and lung cancer [26], but later discovered two at-risk haplotypes and three protective haplotypes in chromosomal region 19q13.2–3 in a small Chinese population, reflecting the more comprehensive nature of the haplotype approach [44]. Recent studies suggest that *ERCC1* haplotypes may also have functional significance. An *in vitro* study by Zhao et al. showed that a specific *ERCC1* haplotype is associated with higher median levels of BPDE-DNA adducts in cultured primary lymphocytes from healthy Caucasians [45]. In a study by Woelfelschneider et al., an *ERCC1* haplotype was associated with lower mRNA levels in lymphocytes from Caucasian prostate cancer patients [46]. Taken together, these findings demonstrate the need to further examine *ERCC1* haplotypes for association with tobacco-induced cancers.

We found that *ERCC1* haplotypes do not modulate lung cancer risk in our Caucasian population. Previously, Ma et al. identified an *ERCC1* haplotype that was protective against lung cancer risk in a Chinese population [30], and another recent study also found an association between *ERCC1* haplotypes and lung cancer risk in a Chinese population [47]. In the latter study, haplotypes containing both –433C and 262G were found to be associated with higher risk for lung cancer, and the risk was higher in individuals that smoked more. In luciferase reporter assays, haplotypes containing the 262G allele displayed significantly lower transcriptional activity as well as lower protein binding measured by EMSA [47]. The 262G allele is much more frequent in Caucasians (92.5%) as compared to Chinese (55.6%). If the 262G>T is the functional SNP, then the 262T allele may indeed exert a protective effect against lung cancer in Caucasians as well as Chinese. However, since the 262T allele has a frequency of only 7.5% in Caucasians, compared to 44.4% in Chinese, this could not be adequately assessed in the Caucasian population described in the present study given the sample size. Moreover, variability in other genes in the NER pathway may help explain the observed racial difference. A comprehensive genetic approach that implements haplotype analyses of multiple genes in the NER pathway may help resolve the differences between Caucasian and Chinese populations.

Tobacco carcinogens are known to induce numerous cancers within the H&N. We present for the first time that *ERCC1* haplotypes are not associated with H&N cancer susceptibility in Caucasians. Low expression of *ERCC1* has been associated with greater risk of squamous cell carcinoma of the head and neck [15]. For this reason, we included a SNP in the promoter region of *ERCC1* in our genetic analysis of this locus. Because the promoter of the *ERCC1* gene lies in a different haplotype block than the coding region, this SNP was necessary to capture any alleles that contribute to expression variability. Even with this SNP included, no association was found between any of the *ERCC1* haplotypes and H&N cancer risk.

This study was adequately powered to detect moderate associations, but sample size was a limiting factor for rare alleles that may be weakly associated with cancer risk. A larger sample size, especially for the H&N cases, would permit detection of such events. Another potential limitation of the study design is that the catchment area for cases was from H. Lee Moffitt Cancer Center oncology clinics, whereas recruited controls were community-based from an affiliated cancer screening clinic. Additionally, patients were enrolled at a single cancer center, and these findings were not validated at an independent center. A particular strength of this study includes a high participation rate for cases (99%) and controls (97%), which limits selection bias [48]. In conclusion, we comprehensively evaluated common genetic variation in the *ERCC1* gene and found no association with lung or H&N cancer risk in Caucasians. Mice with homozygous *ERCC1* mutations exhibit severe aneuploidy within weeks after birth and die before weaning [49]. Given the lack of non-synonymous SNPs in the *ERCC1* coding region, it is possible that the function of this gene is vital and that even minor functional alterations are incompatible with life.

Abbreviations

NER	nucleotide excision repair
ERCC1	excision repair cross-complementation group 1
H&N	head and neck
SNP	single nucleotide polymorphism
XP	xeroderma pigmentosum
LD	linkage disequilibrium
MAF	minor allele frequency
EM	expectation maximization

Acknowledgments

These studies were supported by Public Health Service grants K99-CA 131477 (Gallagher), P01-CA 68384 (Lazarus), R01-DE13158 (Lazarus), K07-CA 104231 (Muscat), from the National Institutes of Health and PA-DOH 4100038714 (Lazarus and Muscat) from the Pennsylvania Department of Health. This project was also funded, in part, under a grant with the Pennsylvania Department of Health using Tobacco Settlement Funds (Spratt). We thank the Functional Genomics Core Facility at the Penn State University College of Medicine for real-time PCR services and Diane McCloskey for editorial assistance.

References

1. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst.* 1999; 91(14):1194–1210. [PubMed: 10413421]

2. Kietthubthaw S, Sriplung H, Au WW, Ishida T. Polymorphism in DNA repair genes and oral squamous cell carcinoma in Thailand. *Int J Hyg Environ Health*. 2006; 209(1):21–29. [PubMed: 16373199]
3. Shen H, Spitz MR, Qiao Y, Guo Z, Wang LE, Bosken CH, et al. Smoking, DNA repair capacity and risk of nonsmall cell lung cancer. *Int J Cancer*. 2003; 107(1):84–88. [PubMed: 12925960]
4. Bessho T, Sancar A, Thompson LH, Thelen MP. Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *J Biol Chem*. 1997; 272(6):3833–3837. [PubMed: 9013642]
5. Araujo SJ, Wood RD. Protein complexes in nucleotide excision repair. *Mutat Res*. 1999; 435(1):23–33. [PubMed: 10526214]
6. Tsodikov OV, Enzlin JH, Scharer OD, Ellenberger T. Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proc Natl Acad Sci U S A*. 2005; 102(32):11236–11241. [PubMed: 16076955]
7. Bryant HE, Ying S, Helleday T. Homologous recombination is involved in repair of chromium-induced DNA damage in mammalian cells. *Mutat Res*. 2006; 599:116. [PubMed: 16564059]
8. Mu D, Bessho T, Nechev LV, Chen DJ, Harris TM, Hearst JE, et al. DNA interstrand cross-links induce futile repair synthesis in mammalian cell extracts. *Mol Cell Biol*. 2000; 20(7):2446–2454. [PubMed: 10713168]
9. Bessho T, Mu D, Sancar A. Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. *Mol Cell Biol*. 1997; 17(12):6822–6830. [PubMed: 9372913]
10. Kiyohara C, Yoshimasu K. Genetic polymorphisms in the nucleotide excision repair pathway and lung cancer risk: a meta-analysis. *Int J Med Sci*. 2007; 4(2):59–71. [PubMed: 17299578]
11. Zienoldindin S, Campa D, Lind H, Ryberg D, Skaug V, Stangeland L, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. *Carcinogenesis*. 2006; 27(3):560–567. [PubMed: 16195237]
12. Zhou W, Liu G, Park S, Wang Z, Wain JC, Lynch TJ, et al. Gene-smoking interaction associations for the ERCC1 polymorphisms in the risk of lung cancer. *Cancer Epidemiol Biomarkers Prev*. 2005; 14(2):491–496. [PubMed: 15734977]
13. Weiss JM, Weiss NS, Ulrich CM, Doherty JA, Voigt LF, Chen C. Interindividual variation in nucleotide excision repair genes and risk of endometrial cancer. *Cancer Epidemiol Biomarkers Prev*. 2005; 14(11 Pt 1):2524–2530. [PubMed: 16284373]
14. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2002; 11(12):1513–1530. [PubMed: 12496039]
15. Cheng L, Sturgis EM, Eicher SA, Spitz MR, Wei Q. Expression of nucleotide excision repair genes and the risk for squamous cell carcinoma of the head and neck. *Cancer*. 2002; 94(2):393–397. [PubMed: 11900225]
16. Winsey SL, Haldar NA, Marsh HP, Bunce M, Marshall SE, Harris AL, et al. A variant within the DNA repair gene XRCC3 is associated with the development of melanoma skin cancer. *Cancer Res*. 2000; 60(20):5612–5616. [PubMed: 11059748]
17. Chen P, Wiencke J, Aldape K, Kesler-Diaz A, Miike R, Kelsey K, et al. Association of an ERCC1 polymorphism with adult-onset glioma. *Cancer Epidemiol Biomarkers Prev*. 2000; 9(8):843–847. [PubMed: 10952103]
18. Cheng L, Spitz MR, Hong WK, Wei Q. Reduced expression levels of nucleotide excision repair genes in lung cancer: a case-control analysis. *Carcinogenesis*. 2000; 21(8):1527–1530. [PubMed: 10910954]
19. Reed E, Yu JJ, Davies A, Gannon J, Armentrout SL. Clear cell tumors have higher mRNA levels of ERCC1 and XPB than other histological types of epithelial ovarian cancer. *Clin Cancer Res*. 2003; 9(14):5299–5305. [PubMed: 14614013]
20. Rosell R, Cecere F, Santarpia M, Reguart N, Taron M. Predicting the outcome of chemotherapy for lung cancer. *Curr Opin Pharmacol*. 2006; 6(4):323–331. [PubMed: 16765644]
21. Rosell R, Taron M, Barnadas A, Scagliotti G, Sarries C, Roig B. Nucleotide excision repair pathways involved in Cisplatin resistance in non-small-cell lung cancer. *Cancer Control*. 2003; 10(4):297–305. [PubMed: 12915808]

22. Shirota Y, Stoecklacher J, Brabender J, Xiong YP, Uetake H, Danenberg KD, et al. ERCC1 and thymidylate synthase mRNA levels predict survival for colorectal cancer patients receiving combination oxaliplatin and fluorouracil chemotherapy. *J Clin Oncol*. 2001; 19(23):4298–4304. [PubMed: 11731512]
23. Dabholkar M, Vionnet J, Bostick-Bruton F, Yu JJ, Reed E. Messenger RNA levels of XPAC and ERCC1 in ovarian cancer tissue correlate with response to platinum-based chemotherapy. *The Journal of Clinical Investigation*. 1994; 94(2):703–708. [PubMed: 8040325]
24. Dabholkar M, Bostick-Bruton F, Weber C, Bohr VA, Egwuagu C, Reed E. ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients. *J Natl Cancer Inst*. 1992; 84(19):1512–1517. [PubMed: 1433335]
25. Vogel U, Laros I, Jacobsen NR, Thomsen BL, Bak H, Olsen A, et al. Two regions in chromosome 19q13.2-3 are associated with risk of lung cancer. *Mutat Res*. 2004; 546(1–2):65–74. [PubMed: 14757194]
26. Yin J, Vogel U, Guo L, Ma Y, Wang H. Lack of association between DNA repair gene ERCC1 polymorphism and risk of lung cancer in a Chinese population. *Cancer Genet Cytogenet*. 2006; 164(1):66–70. [PubMed: 16364765]
27. Li Y, Gu S, Wu Q, Li Y, Fu X, Mao Y, et al. No association of ERCC1 C8092A and T19007C polymorphisms to cancer risk: a meta-analysis. *Eur J Hum Genet*. 2007; 15(9):967–973. [PubMed: 17522621]
28. Sturgis EM, Dahlstrom KR, Spitz MR, Wei Q. DNA repair gene ERCC1 and ERCC2/XPB polymorphisms and risk of squamous cell carcinoma of the head and neck. *Arch Otolaryngol Head Neck Surg*. 2002; 128(9):1084–1088. [PubMed: 12220217]
29. Sugimura T, Kumimoto H, Tohrai I, Fukui T, Matsuo K, Tsurusako S, et al. Gene-environment interaction involved in oral carcinogenesis: molecular epidemiological study for metabolic and DNA repair gene polymorphisms. *J Oral Pathol Med*. 2006; 35(1):11–18. [PubMed: 16393248]
30. Ma H, Xu L, Yuan J, Shao M, Hu Z, Wang F, et al. Tagging single nucleotide polymorphisms in excision repair cross-complementing group 1 (ERCC1) and risk of primary lung cancer in a Chinese population. *Pharmacogenet Genomics*. 2007; 17(6):417–423. [PubMed: 17502833]
31. Zang EA, Wynder EL. Smoking trends in the United States between 1969 and 1995 based on patients hospitalized with non-smoking-related diseases. *Prev Med*. 1998; 27(6):854–861. [PubMed: 9922068]
32. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature*. 2005; 437(7063):1299–1320. [PubMed: 16255080]
33. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–265. [PubMed: 15297300]
34. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet*. 2002; 70(2):425–434. [PubMed: 11791212]
35. Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials*. 1990; 11(2):116–128. [PubMed: 2161310]
36. Belinsky SA, Klinge DM, Stidley CA, Issa JP, Herman JG, March TH, et al. Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer Res*. 2003; 63(21):7089–7093. [PubMed: 14612500]
37. el-Bayoumy K, Upadhyaya P, Desai DH, Amin S, Hoffmann D, Wynder EL. Effects of 1,4-phenylenebis(methylene)selenocyanate, phenethyl isothiocyanate, indole-3-carbinol, and d-limonene individually and in combination on the tumorigenicity of the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in A/J mouse lung. *Anticancer Res*. 1996; 16(5A):2709–2712. [PubMed: 8917375]
38. Gunning WT, Kramer PM, Lubet RA, Steele VE, End DW, Wouters W, et al. Chemoprevention of benzo(a)pyrene-induced lung tumors in mice by the farnesyltransferase inhibitor R115777. *Clin Cancer Res*. 2003; 9(5):1927–1930. [PubMed: 12738751]
39. Keith RL, Miller YE. Lung cancer: genetics of risk and advances in chemoprevention. *Curr Opin Pulm Med*. 2005; 11(4):265–271. [PubMed: 15928489]

40. Hansen RD, Sorensen M, Tjonneland A, Overvad K, Wallin H, Raaschou-Nielsen O, et al. A haplotype of polymorphisms in ASE-1, RAI and ERCC1 and the effects of tobacco smoking and alcohol consumption on risk of colorectal cancer: a Danish prospective case-cohort study. *BMC Cancer*. 2008; 8:54. [PubMed: 18289367]
41. Nexø BA, Vogel U, Olsen A, Ketelsen T, Bukowy Z, Thomsen BL, et al. A specific haplotype of single nucleotide polymorphisms on chromosome 19q13.2-3 encompassing the gene RAI is indicative of post-menopausal breast cancer before age 55. *Carcinogenesis*. 2003; 24(5):899–904. [PubMed: 12771034]
42. Vogel U, Sorensen M, Hansen RD, Tjonneland A, Overvad K, Wallin H, et al. Gene-environment interactions between smoking and a haplotype of RAI, ASE-1 and ERCC1 polymorphisms among women in relation to risk of lung cancer in a population-based study. *Cancer Lett*. 2007; 247(1): 159–165. [PubMed: 16690207]
43. Schierup MH, Mailund T, Li H, Wang J, Tjonneland A, Vogel U, et al. Haplotype frequencies in a sub-region of chromosome 19q13.3, related to risk and prognosis of cancer, differ dramatically between ethnic groups. *BMC Med Genet*. 2009; 10:20. [PubMed: 19257887]
44. Yin J, Vogel U, Ma Y, Qi R, Wang H. Haplotypes of nine single nucleotide polymorphisms on chromosome 19q13.2-3 associated with susceptibility of lung cancer in a Chinese population. *Mutat Res*. 2008; 641(1–2):12–18. [PubMed: 18358500]
45. Zhao H, Wang LE, Li D, Chamberlain RM, Sturgis EM, Wei Q. Genotypes and haplotypes of ERCC1 and ERCC2/XPD genes predict levels of benzo[a]pyrene diol epoxide-induced DNA adducts in cultured primary lymphocytes from healthy individuals: a genotype-phenotype correlation analysis. *Carcinogenesis*. 2008; 29(8):1560–1566. [PubMed: 18635523]
46. Woelfelschneider A, Popanda O, Lilla C, Linseisen J, Mayer C, Celebi O, et al. A distinct ERCC1 haplotype is associated with mRNA expression levels in prostate cancer patients. *Carcinogenesis*. 2008; 29(9):1758–1764. [PubMed: 18332046]
47. Yu D, Zhang X, Liu J, Yuan P, Tan W, Guo Y, et al. Characterization of functional excision repair cross-complementation group 1 variants and their association with lung cancer risk and prognosis. *Clin Cancer Res*. 2008; 14(9):2878–2886. [PubMed: 18451256]
48. Austin H, Hill HA, Flanders WD, Greenberg RS. Limitations in the application of case-control methodology. *Epidemiol Rev*. 1994; 16(1):65–76. [PubMed: 7925729]
49. McWhir J, Selfridge J, Harrison DJ, Squires S, Melton DW. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat Genet*. 1993; 5(3):217–224. [PubMed: 8275084]

Figure 1.

LD structure of *ERCCI* and 10-kb promoter. The boundaries of the coding region block (block 1) and promoter block (block 2) are defined by the solid black triangular lines. *ERCCI* HapMap genotypes from Caucasians (CEPH) were downloaded and LD was determined using Haploview. D' values are displayed in the squares (empty squares have pairwise $D'=1.0$). Red squares show high pairwise LD, while light red and white squares have increasingly lower LD. Tag-SNPs are denoted by rectangular boxes, with three tagSNPs (rs3212948, rs3212955, and rs735482) in the coding region block and one tagSNP (rs1319052) in the promoter block.

Table 1

Age, gender, BMI, smoking status, education level, and case histology of 452 lung cancer cases and 790 controls.

	Lung cases n (%)	Controls n (%)	p-value
Mean Age ^a	64.2 ± 9.8	58.4 ± 10.4	<0.01
Women	206 (46)	365 (46)	0.83
Mean BMI ^a	26.9 ± 5.1	27.2 ± 4.9	0.30
Smoking Status			<0.01
Never Smokers	39 (9)	305 (39)	
Former Smokers	229 (51)	328 (42)	
Current Smokers	184 (41)	157 (20)	
Pack-Years ^a	56.4 ± 38.6	23.8 ± 30.8	<0.01
Education Level			<0.01
No Degree	68 (15)	35 (4)	
High School Degree	289 (64)	447 (57)	
College Degree	63 (14)	186 (24)	
Graduate Degree	32 (7)	122 (15)	
Histology			
Adenocarcinoma	172 (38)		
Squamous Cell Carcinoma	109 (24)		
Non-small Cell Carcinoma	84 (19)		
Small-Cell Carcinoma	41 (9)		
Large-Cell Carcinoma	29 (6)		
Mixed Histology/Other	17 (4)		

^a mean ± standard deviation for continuous variables

Table 2

Age, gender, BMI, smoking status, and education level of 175 head and neck cancer cases and 790 controls.

	H&N cases n (%)	Controls n (%)	p-value
Mean Age ^a	58.5 ± 11.4	58.4 ± 10.4	0.95
Women	46 (26)	365 (46)	<0.01
Mean BMI ^a	26.7 ± 5.5	27.2 ± 4.9	0.29
Smoking Status			<0.01
Never Smokers	44 (25)	305 (39)	
Former Smokers	56 (32)	328 (42)	
Current Smokers	75 (43)	157 (20)	
Pack-Years ^a	48.4 ± 42.2	23.8 ± 30.8	<0.01
Education Level			<0.01
No Degree	23 (13)	35 (4)	
High School Degree	112 (64)	447 (57)	
College Degree	27 (15)	186 (24)	
Graduate Degree	12 (7)	122 (15)	

^a mean ± standard deviation for continuous variables

Table 3Distribution of *ERCC1* SNPs and lung cancer risk.

Genotype	Cases, n (%)	Controls, n (%)	OR (95% CI) ^a
rs1319052 (G>A)			
GG	169 (40.6)	271 (35.9)	1.00 (reference)
GA	186 (44.7)	361 (47.8)	0.82 (0.61–1.11)
AA	61 (14.7)	123 (16.3)	0.99 (0.66–1.48)
rs3212948 (C>G)			
CC	184 (43.1)	276 (37.4)	1.00 (reference)
CG	184 (43.1)	348 (47.2)	0.84 (0.62–1.12)
GG	59 (13.8)	113 (15.3)	0.91 (0.60–1.37)
rs3212955 (A>G)			
AA	245 (58.9)	428 (56.5)	1.00 (reference)
AG	148 (35.6)	288 (38.0)	0.97 (0.73–1.28)
GG	23 (5.5)	42 (5.5)	0.95 (0.52–1.72)
rs735482 (A>C)			
AA	289 (72.6)	523 (73.0)	1.00 (reference)
AC	102 (25.6)	175 (24.4)	1.01 (0.73–1.38)
CC	7 (1.8)	18 (2.5)	0.89 (0.33–2.40)

^aEstimated by logistic regression, adjusting for age, gender, and pack-years

Table 4Distribution of *ERCC1* SNPs and head and neck cancer risk.

Genotype	Cases, n (%)	Controls, n (%)	OR (95% CI) ^a
rs1319052 (G>A)			
GG	57 (32.9)	271 (35.9)	1.00 (reference)
GA	95 (54.9)	361 (47.8)	1.29 (0.88–1.89)
AA	21 (12.1)	123 (16.3)	0.90 (0.51–1.59)
rs3212948 (C>G)			
CC	64 (37.2)	276 (37.4)	1.00 (reference)
CG	88 (51.2)	348 (47.2)	1.14 (0.78–1.67)
GG	20 (11.6)	113 (15.3)	0.82 (0.46–1.45)
rs3212955 (A>G)			
AA	92 (52.9)	428 (56.5)	1.00 (reference)
AG	70 (40.2)	288 (38.0)	1.13 (0.78–1.63)
GG	12 (6.9)	42 (5.5)	1.36 (0.67–2.75)
rs735482 (A>C)			
AA	122 (76.3)	523 (73.0)	1.00 (reference)
AC	37 (23.1)	175 (24.4)	0.88 (0.58–1.35)
CC	1 (0.6)	18 (2.5)	0.32 (0.04–2.49)

^aEstimated by logistic regression, adjusting for age, gender, and pack-years

Table 5

Distribution of *ERCC1* haplotypes and lung cancer risk.

	SNP ^a			Lung cases		Controls		Global p-value = 0.434
	I	II	III	% (n = 452)		% (n = 790)		p-value ^b
Haplotype 1	C	A	A	60.0		58.2		0.727
Haplotype 2	G	G	A	23.7		24.7		0.811
Haplotype 3	G	A	C	10.0		12.0		0.227
Haplotype 4	C	A	C	4.6		2.9		0.141

^a I [rs3212948, C>G]; II [rs3212955, A>G]; III [rs735482, A>C]

^b p-values were adjusted for age, gender, and pack-years

Table 6

Distribution of *ERCC1* haplotypes and head and neck cancer risk.

	SNP ^a			H&N cases		Controls		Global p-value = 0.534
	I	II	III	% (n = 175)		% (n = 790)		p-value ^b
Haplotype 1	C	A	A	60.3	58.2			0.548
Haplotype 2	G	G	A	25.6	24.7			0.615
Haplotype 3	G	A	C	10.1	12.0			0.273
Haplotype 4	C	A	C	1.9	2.9			0.498

^a I [rs3212948, C>G]; II [rs3212955, A>G]; III [rs735482, A>C]

^b p-values were adjusted for age, gender, and pack-years