Int. J. Cancer: **122,** 1630–1638 (2008)

A case-control study of the association of the polymorphisms and haplotypes of DNA ligase I with lung and upper-aerodigestive-tract cancers

Yuan-Chin Amy Lee[1], Hal Morgenstern[2,3], Sander Greenland[1,4], Donald P. Tashkin[5], Jeanette Papp[6], Janet Sinsheimer[6], Wei Cao[1], Mia Hashibe[7], Nai-Chieh Y. You[1], Jenny T. Mao[5], Wendy Cozen[8], Thomas M. Mack[8] and Zuo-Feng Zhang[1]*

[1]*Department of Epidemiology, University of California at Los Angeles School of Public Health, Los Angeles, CA*
[2]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI*
[3]*Department of Environmental Health Sciences, University of Michigan School of Public Health, Ann Arbor, MI*
[4]*Department of Statistics, University of California at Los Angeles, Los Angeles, CA*
[5]*Division of Pulmonary and Critical Care Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA*
[6]*Department of Human Genetics, University of California at Los Angeles, Los Angeles, CA*
[7]*International Agency for Research on Cancer, Lyon, France*
[8]*Department of Preventive Medicine, Keck School of Medicine at University of Southern California, Los Angeles, CA*

Tobacco smoking is a major risk factor for lung and upper-aerodigestive-tract (UADT) cancers. One possible mechanism for the associations may be through DNA damage pathways. DNA Ligase I (*LIG1*) is a DNA repair gene involved in both the nucleotide excision repair (NER) and the base excision repair (BER) pathways. We examined the association of 4 *LIG1* polymorphisms with lung and UADT cancers, and their potential interactions with smoking in a population-based case-control study in Los Angeles County. We performed genotyping using the SNPlex method from Applied Biosystems. Logistic regression analyses of 551 lung cancer cases, 489 UADT cancer cases and 948 controls showed the expected associations of tobacco smoking with lung and UADT cancers and new associations between the *LIG1* haplotypes and these cancers. For lung cancer, when compared to the most common haplotype (rs20581-rs20580-rs20579-rs439132 = T-C-C-A), the adjusted odds ratio (OR) is 1.2 (95% confidence limits (CL) = 0.95, 1.5) for the CACA haplotype, 1.4 (1.0, 1.9) for the CATA haplotype and 1.8 (1.1, 2.8) for the CCCG haplotype, after controlling for age, gender, race/ethnicity, education and tobacco smoking. We observed weaker associations between the *LIG1* haplotypes and UADT cancers. Our findings suggest the *LIG1* haplotypes may affect the risk of lung and UADT cancers.

Key words: LIG1; polymorphisms; haplotypes; lung cancer; upper-aerodigestive-tract cancers

Lung cancer patients have the poorest survival rates among cancer patients. The 5-year survival rate of lung cancer is only 11% worldwide.[1] It is also the most common cancer in the world and the third most common cancer in the United States.[1–3] In the United States, lung cancer incidence ranks number three, while lung cancer mortality ranks number one in 2006.[3] Although several risk factors for lung cancer such as tobacco smoking have been identified,[4] little is known about the role of genetic susceptibility and its potential interactions with environmental factors. Studies have examined associations between lung cancer and the genetic polymorphisms of DNA repair genes; however, there are no conclusive results on potential interactions with environmental factors.

Upper-aerodigestive-tract (UADT) cancers of the oral cavity, pharynx, larynx and esophagus accounts for 10% of new cases and 11% of cancer deaths worldwide in 2002.[1] In the United States, the estimated incidence and death rates are high in 2006.[3] Oral/pharyngeal, laryngeal and esophageal cancers show 5-year survival rates of ~59, 64 and 16%, respectively, during 1995–2000.[5] Tobacco and alcohol exposures have been well documented as risk factors for head and neck squamous cell carcinomas.[6–8] However, results of previous studies on the potential associations between UADT cancers and the genetic polymorphisms of DNA repair genes and their possible interactions with environmental factors have been inconclusive.[9,10]

DNA is vulnerable to injury due to endogenous and/or exogenous exposures[11] such as exposure to carcinogens in tobacco smoke. DNA repair systems play an important role in maintaining the integrity of the genome. Deficiencies in this defense system are thought to contribute to the development of cancer. Based on a review of epidemiologic studies, Berwick and Vineis[12] suggested that reduced DNA repair capacity was associated with increased risk of cancers, with odds ratios in the range of 1.4–75, with the majority of values between 2 and 10. In addition, genetic variation in DNA repair genes may influence DNA repair capacity; consequently, DNA repair genetic polymorphisms may affect cancer susceptibility.

DNA ligase I (*LIG1*) is a 102-kDa nuclear enzyme, which is involved in joining Okazaki fragments during DNA replication and in sealing single-strand breaks.[13–15] It plays a role in both the nucleotide excision repair (NER) and the long-patch base-excision repair (BER) pathways.[16,17] NER is an important DNA repair pathway against tobacco-related carcinogenesis.[18] A total of 31 SNPs for *LIG1* were identified in SNP500 database (http://snp500cancer.nci.nih.gov/snplist.cfm). The polymorphic sites selected are shown in Figure 1. The T → C synonymous SNP on exon 25 (rs20581), the C → A synonymous SNP on exon 7 (rs20580), C → T synonymous SNP at 5′ UTR on exon 2 (rs20579) and A → G SNP on intron 1 (rs439132) do not result in any alteration of amino acid[19] and their functional variation is unknown.[20] They were selected because the *LIG1* gene plays a very important role in the BER pathway, and their rare allele frequencies are greater or equal to 5%. The D' between rs20581 and rs20580 is 1.0; the D' between rs20580 and rs20579 is 0.842; the D' between rs20581 and rs20579 is 1.0. These polymorphisms might affect the transcription or translation of LIG1 protein for DNA repair and, in turn, may influence DNA repair capacity. The effect would thus take place after damage has occurred during carcinogenesis. These 4 SNPs may be associated with cancer risks separately or jointly. Therefore, the haplotypes of *LIG1* may play a very important role.

We are aware of only one epidemiologic study that has examined the association between one of the *LIG1* polymorphisms

Abbreviations: BER, base excision repair; CL, confidence limits; *LIG1*, DNA ligase 1; NER, nucleotide excision repair; OR, odds ratio; SNP, single nucleotide polymorphism; UADT cancers, upper-aerodigestive-tract cancers including oral/pharyngeal, laryngeal and esophageal cancers.

Grant sponsor: National Institute of Health; Grant number: CA90833, DA11386, CA77954, CA09142, CA96134, ES 011667, GM53275. Grant sponsor: UCLA Jonsson Comprehensive Cancer Center (Alper Research Program for Environmental Genomics).

***Correspondence to:** Department of Epidemiology, University of California at Los Angeles School of Public Health, 71-225 CHS, Box 951772, 650 Charles E Young Drive, South, Los Angeles, CA, 90095-1772, USA. Fax: +310-206-6039. E-mail: zfzhang@ucla.edu

Received 7 March 2007; Accepted after revision 18 September 2007
DOI 10.1002/ijc.23274
Published online 4 December 2007 in Wiley InterScience (www.interscience.wiley.com).

(rs20580) and lung cancer.[20] That study, however, did not examine the associations of the other *LIG1* SNPs, haplotypes and lung cancer, and the joint effects of the *LIG1* SNPs with tobacco smoking. We know of no published studies that have explored the associations between these *LIG1* SNPs, haplotypes and lung and UADT cancers. The objectives of the present study were to evaluate the associations between the DNA repair genetic polymorphisms and haplotypes of DNA Ligase I (*LIG1*) and lung and UADT cancer risks in the residential population of Los Angeles County, and to investigate possible interactions between the *LIG1* SNPs and haplotypes and environmental factors, particularly tobacco smoking.

Material and methods

Study design and study participants

We used questionnaire data and buccal cell samples from a population-based case-control study, including 611 incident cases of lung cancer, 601 incident cases of UADT cancer (403 oral/pharyngeal, 90 laryngeal and 108 esophageal) and 1,040 population controls.[21] Histologically confirmed newly diagnosed cancer cases were obtained by the rapid ascertainment system of the Cancer Surveillance Program for Los Angeles County, which is administered by the Keck School of Medicine and Norris Comprehensive Cancer Center at the University of Southern California. Over 95% of cancer reports were histologically verified; the remainder was verified by MRI, CT scan or other diagnostic methods. Controls did not have a history of oral/pharyngeal, laryngeal, esophageal or lung cancer and were identified using a formal algorithm that enumerated households in the neighborhoods of the cases. Controls were individually matched to cases by age (within 10-year categories), gender and neighborhood.

All participants met the following criteria: (*i*) were residents of Los Angeles County at the time of diagnosis for cases or at the time of recruitment for controls, (*ii*) were 18–62 years of age during the study period of 1999–2004 and (*iii*) spoke either English or Spanish, or had translators available at the time of interview. Additional details regarding recruitment and other exclusion criteria have been described previously.[21,22] Each case and control participated in in-person interviews consisting of standardized questionnaires administered by specially trained interviewers. After the interview, buccal swabs were collected for subsequent DNA extraction.

Of the lung cancer and UADT cancer cases, respectively, 57 and 49% were interviewed within 3 months of diagnosis, 32 and 38% 3–6 months after diagnosis, and 11 and 12% more than 6 months after diagnosis. The response rates for those who were eligible to be in the study were 68, 39, 49, 40 and 34% for controls, lung cancer, oral/pharyngeal cancer, laryngeal cancer, and esophageal cancer cases, respectively. Furthermore, the response rates for those who completed interviews to provide the biospecimen were 89, 89, 68, 88 and 90% for controls, lung cancer, oral/pharyngeal cancer, laryngeal cancer and esophageal cancer cases, respectively. The protocol was approved by the Institutional Review Boards of the University of California, Los Angeles and the University of Southern California. Informed consent was obtained from every study subject.

Buccal cell samples

After participants were interviewed, interviewers collected ~25–50 ml of buccal cells with Scope™ mouthwash solution in tubes. All buccal cell samples were transported as soon as possible to the School of Public Health, University of California, Los Angeles where samples were stored at 4°C in the molecular epidemiology laboratory refrigerator. The buccal cell samples were centrifuged to separate the buccal cells and mouthwash solution. High molecular weight DNA was isolated by an altered phenol-chloroform method.[23] DNA purity and concentration were determined by a spectrophotometric measurement of absorbance at 260/280 nm.

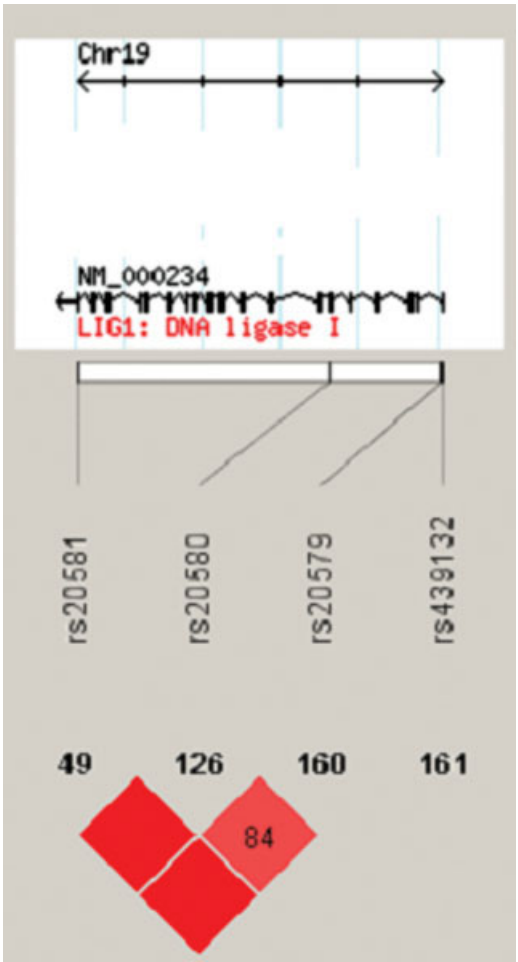

Figure 1 – The selected polymorphic sites of LIG1. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Genotyping assays

Genotypes of *LIG1* were determined using the Applied Biosystems SNPlex™ assay.[24] This method assays SNP genotypes in 48 SNP PCR multiplexes using an oligonucleotide ligation process. Each SNP allele was ligated to an allele-specific oligonucleotide, then hybridized with a Zipchute™ probe with a mobility modifier and a fluorescent label, allowing the probe to be separated and detected by capillary electrophoresis. Detection was performed on an Applied Biosystems 3730 DNA Analyzer, and data interpretation was performed with the Applied Biosystems Genemapper software v4.0, which uses a clustering algorithm with stringent proprietary quality checks to call genotypes. SNPlex genotyping was performed by the UCLA Genotyping Core. All automatic calls by the software were evaluated by Core technicians. Any SNPs with less than 90% of the sample auto-called by the software were either rescored manually or discarded if clustering confidence was low. Samples were run with 2 positive controls of CEPH individual 1347-2 and 2 "no template", negative controls per 96-well plate. Any run in which the positive controls failed to cluster was discarded. Only genotypes with a Genemapper Quality Score of greater than 95% were passed. On the DNA samples in this study, the call rate was 96%, the reproducibility rate was 97.8% and the concordance rate was 99.8%. Reproducibility was determined by comparing blinded replicates on plates and by rerunning entire plates. The reproducibility reported was calculated over 10,000 duplicate genotypes (3% of the entire sample), including not only blinded replicates on the same plate or the

TABLE 1 – DEMOGRAPHIC CHARACTERISTICS OF LUNG, UADT CANCER CASES AND POPULATION CONTROLS STRATIFIED BY THOSE WITH AND WITHOUT DNA SPECIMENS

Variables	Lung cancer cases (n = 611), N (%)		Controls (n = 1029), N (%)		UADT cancer cases (n = 601), N (%)		Controls (n = 1039), N (%)	
	With DNA (551)	Without DNA (60)	With DNA (939)	Without DNA (90)	With DNA (489)	Without DNA (112)	With DNA (948)	Without DNA (91)
Mean age (years), (age range)	52.2 (32–59)	52.3 (41–59)	50.2 (29–62)	49.5 (30–60)	50.6 (23–59)	49.6 (20–59)	49.9 (17–62)	49.2 (24–60)
	p-value = 0.90		p-value = 0.37		p-value = 0.21		p-value = 0.38	
Age (years)								
<35	4 (0.73)	0 (0.0)	38 (4.1)	3 (3.3)	22 (5.0)	10 (6.0)	47 (5.0)	4 (3.0)
35–44	51 (9.3)	6 (10)	155 (17)	16 (18)	54 (13)	23 (13)	155 (16)	22 (16)
45–54	273 (50)	28 (47)	452 (48)	47 (52)	184 (43)	83 (46)	452 (48)	70 (52)
≥55	223 (40)	26 (43)	294 (31)	24 (27)	163 (39)	62 (35)	294 (31)	40 (29)
	p-value = 0.88		p-value = 0.79		p-value = 0.85		p-value = 0.64	
Race/Ethnicity								
Caucasian	320 (58)	39 (65)	583 (62)	45 (50)	286 (59)	92 (52)	588 (62)	46 (51)
Hispanic	50 (9.1)	3 (5.0)	140 (15)	8 (8.9)	60 (12)	22 (12)	141 (15)	8 (8.8)
African American	90 (16)	6 (10)	79 (8.4)	23 (25)	49 (10)	28 (16)	79 (8.3)	23 (25)
Asian American	59 (11)	11 (18)	53 (5.7)	7 (7.8)	50 (10)	20 (11)	55 (5.8)	7 (7.7)
Other	31 (5.6)	1 (1.7)	83 (8.9)	7 (7.8)	43 (8.8)	15 (9.0)	84 (8.9)	7 (7.7)
Missing	1	0	1	0	1	1	1	0
	p-value = 0.13		p-value < 0.0001		p-value = 0.09		p-value < 0.0001	
Gender								
Male	278 (50)	25 (42)	575 (61)	41 (46)	380 (78)	74 (66)	582 (61)	41 (45)
Female	273 (50)	35 (58)	364 (39)	49 (54)	109 (22)	38 (34)	366 (39)	50 (55)
	p-value = 0.20		p-value = 0.004		p-value = 0.010		p-value = 0.002	
Education Level								
≤12 years	248 (45)	17 (28)	276 (29)	23 (26)	219 (45)	54 (48)	276 (29)	23 (25)
>12 to ≤16 years	248 (45)	27 (45)	422 (45)	49 (54)	211 (43)	48 (43)	431 (46)	50 (55)
>16 years	55 (10)	16 (27)	240 (26)	18 (20)	59 (12)	10 (8.9)	240 (25)	18 (20)
Missing	0	0	1	0	0	0	1	0
	p-value = 0.0003		p-value = 0.22		p-value = 0.60		p-value = 0.22	

TABLE II – THE ASSOCIATIONS BETWEEN TOBACCO SMOKING, LUNG CANCER, AND UADT CANCER

Tobacco smoking	Lung cancer			UADT cancers		
	Cases *N*	Controls *N*	Adjusted OR[1] (95% CL)	Cases *N*	Controls *N*	Adjusted OR[2] (95% CL)
Smoking status						
Never	110	484	1.0	182	491	1.0
Ever	501	545	4.1 (3.2, 5.3)	419	548	1.5 (1.2, 1.9)
Former	360	361	4.4 (3.4, 5.7)	314	362	1.8 (1.4, 2.3)
Current	141	184	3.7 (2.7, 5.0)	105	186	0.93 (0.67, 1.3)
			$p_{trend} < 0.0001$			$p_{trend} = 0.45$
Pack-years						
0	110	484	1.0	182	491	1.0
>0–20	102	350	1.4 (0.99, 1.9)	147	353	0.99 (0.76, 1.3)
>20–40	202	136	8.3 (6.0, 12)	146	136	2.3 (1.7, 3.2)
>40	197	58	22 (15, 33)	126	58	4.3 (2.9, 6.4)
			$p_{trend} < 0.0001$			$p_{trend} < 0.0001$
Age started smoking						
Never	110	484	1.0	182	491	1.0
>16	222	317	3.1 (2.3, 4.1)	182	319	1.3 (0.96, 1.6)
≤16	279	227	5.7 (4.3, 7.6)	237	228	1.9 (1.5, 2.5)
			$p_{trend} < 0.0001$			$p_{trend} = 0.22$
Years since quitting smoking						
0	141	184	1.0	105	186	1.0
1–4	241	55	5.4 (3.7, 7.9)	201	56	7.7 (5.1, 12)
>4–12	64	66	1.1 (0.74, 1.7)	37	66	1.2 (0.73, 2.0)
>12	55	240	0.27 (0.19, 0.40)	76	240	0.71 (0.48, 1.0)
			$p_{trend} < 0.0001$			$p_{trend} < 0.0001$

[1]Adjusted for age, gender, ethnicity and education.–[2]Adjusted for age, gender, ethnicity, education and alcohol-years (alcohol drinks/day × years).

same run, but also plates repeated on the entire assay, including all preprocessing steps, as much as 6 months apart. Concordance was determined by Taqman assay on ABI 7900. For concordance tests, we chose one SNP with high call rates, one average SNP and one SNP with low call rates (~60%). These SNPs were genotyped on one 96-well plate with high-quality DNA and high call rates and one plate with low-quality DNA and low call rates. The concordance rate of genotypes called successfully on both SNPlex and Taqman was not affected by SNP quality or DNA quality, although the samples with low call rates on SNPlex also had low call rates on Taqman. All genotypes were checked for Hardy-Weinberg equilibrium and divergence from published allele frequencies. Checks were performed across the entire sample and on individual genotyping runs to identify any possibly bad runs. Any markers, sample sets or genotyping runs that gave suspect results on HWE or on any of our other quality control validation tests were manually rescored.

Statistical analyses

Data analyses were performed with SAS 9 software. All P-values are 2-sided. To increase the precision and power over standard matched analyses, we compared each cancer group (lung and UDAT) with all controls using unconditional logistic regression, including terms for the matching factors, age and sex, which allowed us to include cases with no matched controls.[25] Odds-ratio estimates (ORs) and 95% confidence limits (CLs) were computed with and without adjustment for potential confounders including age, gender, race/ethnicity (white, Hispanic, African American, Asian/Pacific Islander and other), educational level, alcohol drinking and tobacco smoking. To minimize age confounding and to account for age matching, age was controlled in fine categories (<35, 35–36, 37–38, 39–40, 41–42, 43–44, 45–46, 47–48, 49–50, 51–52, 53–54, 55–56, 57–58 and 59–62). Controls who were more than 3 years younger than the youngest case or 3 years older than the oldest case were excluded from the analysis. Educational level (years of schooling) and tobacco smoking (pack-years) were controlled as continuous variables.

We assessed the results for the *LIG1* polymorphisms by treating participants with the homozygote wild-type genotype as the referent. We performed the Pearson (χ^2) test to evaluate Hardy-Weinberg equilibrium among the controls. The haplotypes were reconstructed from our population genotype data, using PHASE v2.[26,27] We treated the most frequent haplotype as the referent. The impact of tobacco smoking on the association of the *LIG1* genotypes and haplotypes was also investigated by stratified analysis.

Results

Lung cancer cases ranged from 32 to 59 years of age, while oral/pharyngeal, laryngeal and esophageal cancer cases ranged from 17 to 59 years of age. Based on the age restriction mentioned previously, eleven controls were excluded from the lung-cancer analyses, and one control was excluded from the UADT-cancer analyses.

Among the lung cancer cases, there were 95 squamous cell carcinoma, 297 adenocarcinoma, 115 large cell lung cancer and 75 small cell lung cancer; among the UADT cancer cases, there were 497 squamous cell carcinomas. Table I shows the comparisons between interviewed cases and controls with and without DNA specimens. There were appreciable differences in race/ethnicity between those with and without DNA specimens among both UADT cancer cases and population controls; African Americans were less likely to provide DNA specimens. Women were less likely than men to provide DNA specimens in all case and control groups, and college-educated cases of lung cancer were less likely to provide DNA specimens than those with less education. The association of DNA sampling with these covariates implies that adjustment for them would be needed.

Table II shows the associations between tobacco smoking and lung and UADT cancer risk. Tobacco smoking (> 40 pack-years *vs*. 0 pack-years) was strongly associated with increased cancer risk with odds ratios of 22 (95% CL = 15, 33) for lung cancer and 4.3 (95% CL = 2.9, 6.4) for UADT cancers after adjusting for age, gender, ethnicity and education (plus alcohol drinking for UADT cancers only). The age at initiation of smoking (≤16-years old and >16-years old *vs*. never started) was also strongly associated with lung cancer (OR = 5.7, 95% CL = 4.3, 7.6 for ≤16-

TABLE III – THE ASSOCIATIONS BETWEEN THE FOUR *LIG1* POLYMORPHISMS AND LUNG CANCER

Genotype of LIG1	Cases	Controls	Crude OR	Adj. OR[1]
Rs20581 (Exon 25)				
T/T	78	142	1.0	1.0
C/T	148	346	0.78 (0.56, 1.1)	0.92 (0.62, 1.4)
C/C	86	155	1.0 (0.69, 1.5)	1.2 (0.74, 1.8)
Stratified by smoking status in pack-years[2]				
0				
T/T	12	62	1.0	1.0
C/T	23	164	0.73 (0.34, 1.5)	0.70 (0.32, 1.5)
C/C	21	71	1.5 (0.70, 3.4)	1.1 (0.49, 2.7)
>0–20				
T/T	7	45	1.0	1.0
C/T	31	116	1.7 (0.71, 4.2)	1.6 (0.62, 3.9)
C/C	15	60	1.6 (0.61, 4.3)	1.5 (0.54, 4.1)
>20				
T/T	59	35	1.0	1.0
C/T	94	66	0.85(0.50, 1.4)	0.81 (0.47, 1.4)
C/C	50	23	1.3 (0.68, 2.5)	1.1 (0.58, 2.2)
Rs20580 (Exon 7)				
C/C	110	202	1.0	1.0
A/C	211	398	0.97 (0.73, 1.3)	1.0 (0.72, 1.4)
A/A	121	179	1.2 (0.90, 1.7)	1.2 (0.85, 1.8)
Stratified by smoking status in pack-years[2]				
0				
C/C	20	97	1.0	1.0
A/C	36	183	0.95 (0.52, 1.7)	0.83 (0.44, 1.6)
A/A	24	80	1.5 (0.75, 2.8)	1.1 (0.55, 2.3)
>0–20				
C/C	14	62	1.0	1.0
A/C	34	139	1.1 (0.54, 2.2)	1.1 (0.52, 2.2)
A/A	21	66	1.4 (0.66, 3.0)	1.5 (0.67, 3.2)
>20				
C/C	76	43	1.0	1.0
A/C	141	76	1.1 (0.66, 1.7)	1.0 (0.64, 1.7)
A/A	76	33	1.3 (0.75, 2.3)	1.2 (0.71, 2.2)
Rs20579 (Exon 2)				
C/C	294	586	1.0	1.0
C/T	118	187	1.3 (0.96, 1.6)	1.3 (0.97, 1.8)
T/T	11	7	3.1 (1.2, 8.2)	2.1 (0.74, 5.9)
Stratified by smoking status in pack-years[2]				
0				
C/C	58	276	1.0	1.0
C/T	15	88	0.81 (0.44, 1.5)	0.84 (0.44, 1.6)
T/T	1	0	–	–
>0–20				
C/C	40	192	1.0	1.0
C/T	23	66	1.7 (0.93, 3.0)	1.7 (0.92, 3.1)
T/T	4	4	4.8 (1.2, 20)	3.8 (0.86, 17)
>20				
C/C	196	118	1.0	1.0
C/T	80	33	1.5 (0.92, 2.3)	1.5 (0.91, 2.4)
T/T	6	3	1.2 (0.30, 4.9)	1.2 (0.30, 5.2)
Rs439132 (Intron 1)				
A/A	326	585	1.0	1.0
A/G	39	54	1.3 (0.84, 2.0)	1.3 (0.81, 2.3)
G/G	6	2	5.4 (1.1, 27)	5.3 (0.92, 30)
Stratified by smoking status in pack-years[2]				
0				
A/A	51	284	1.0	1.0
A/G	10	23	2.4 (1.1, 5.4)	1.7 (0.74, 4.1)
G/G	0	1	–	–
>0–20				
A/A	47	184	1.0	1.0
A/G	5	24	0.82 (0.30, 2.3)	0.78 (0.27, 2.2)
G/G	2	1	7.8 (0.70, 88)	5.2 (0.44, 60)
>20				
A/A	228	117	1.0	1.0
A/G	24	7	1.8 (0.74, 4.2)	1.5 (0.61, 3.8)
G/G	4	0	–	–

[1]Adjusted for gender, age, education, ethnicity and tobacco smoking (pack-years).–[2]Adjusted for gender, age, education and ethnicity.

TABLE IV – THE ASSOCIATIONS BETWEEN THE FOUR *LIG1* POLYMORPHISMS AND UADT CANCER

Genotype of LIG1	Cases	Controls	Crude OR	Adj. OR[1]
Rs20581 (Exon 25)				
T/T	61	143	1.0	1.0
C/T	164	350	1.1 (0.77, 1.6)	1.2 (0.85, 1.8)
C/C	94	157	1.4 (0.95, 2.1)	1.5 (1.0, 2.3)
Stratified by smoking status in pack-years[2]				
0				
T/T	23	62	1.0	1.0
C/T	58	167	0.94 (0.53, 1.6)	0.89 (0.50, 1.6)
C/C	25	72	0.94 (0.48, 1.8)	0.83 (0.42, 1.6)
>0–20				
T/T	9	46	1.0	1.0
C/T	47	117	2.1(0.93, 4.5)	1.7 (0.74, 4.8)
C/C	28	61	2.3 (1.0, 5.5)	2.3 (0.95, 5.4)
>20				
T/T	29	35	1.0	1.0
C/T	59	66	1.1 (0.59, 2.0)	1.2 (0.62, 2.2)
C/C	41	23	2.2 (1.1, 4.4)	2.2 (1.0, 4.7)
Rs20580 (Exon 7)				
C/C	91	203	1.0	1.0
A/C	197	403	1.1 (0.81, 1.5)	1.1 (0.83, 1.6)
A/A	99	180	1.2 (0.87, 1.7)	1.2 (0.83, 1.7)
Stratified by smoking status in pack-years[2]				
0				
C/C	34	98	1.0	1.0
A/C	64	187	0.99 (0.61, 1.6)	0.98 (0.60, 1.6)
A/A	24	80	0.86 (0.47, 1.6)	0.86 (0.47, 1.6)
>0–20				
C/C	16	67	1.0	1.0
A/C	50	140	1.4 (0.73, 2.6)	1.3 (0.68, 2.5)
A/A	26	62	1.5 (0.74, 3.1)	1.5 (0.74, 3.2)
>20				
C/C	41	43	1.0	1.0
A/C	83	76	1.1 (0.68, 1.9)	1.3 (0.73, 2.3)
A/A	49	33	1.6 (0.84, 2.9)	1.6 (0.84, 3.1)
Rs20579 (Exon 2)				
C/C	264	592	1.0	1.0
C/T	108	188	1.3 (0.98, 1.7)	1.3 (1.0, 1.8)
T/T	8	7	2.6 (0.92, 7.1)	2.0 (0.69, 5.8)
Stratified by smoking status in pack-years[2]				
0				
C/C	80	281	1.0	1.0
C/T	35	88	1.4 (0.88, 2.2)	1.5 (0.91, 2.3)
T/T	3	0	–	–
>0–20				
C/C	67	193	1.0	1.0
C/T	24	67	1.0 (0.60, 1.8)	1.1 (0.61, 1.9)
T/T	2	4	1.4 (0.26, 8.0)	1.2 (0.21, 6.9)
>20				
C/C	117	118	1.0	1.0
C/T	49	33	1.5 (0.90, 2.5)	1.5 (0.88, 2.6)
T/T	3	3	1.0 (0.20, 5.1)	0.95 (0.18, 5.0)
Rs439132 (Intron 1)				
A/A	297	590	1.0	1.0
A/G	23	54	0.85 (0.51, 1.4)	0.78 (0.45, 1.4)
G/G	8	2	7.9 (1.7, 38)	5.9 (1.1, 31)
Stratified by smoking status in pack-years[2]				
0				
A/A	100	288	1.0	1.0
A/G	5	23	0.63 (0.23, 1.7)	0.58 (0.21, 1.6)
G/G	1	1	2.9 (0.18, 46)	2.7 (0.16, 45)
>0–20				
A/A	70	185	1.0	1.0
A/G	3	24	0.33 (0.096, 1.1)	0.32 (0.092, 1.1)
G/G	1	1	2.6 (0.16, 43)	3.3 (0.18, 60)
>20				
A/A	127	117	1.0	1.0
A/G	15	7	2.0 (0.78, 5.0)	2.0 (0.74, 5.7)
G/G	6	0	–	–

[1]Adjusted for gender, age, education, ethnicity and tobacco smoking (pack-years).–[2]Adjusted for gender, age, education and ethnicity.

TABLE V – THE ASSOCIATIONS OF THE 4 *LIG1* HAPLOTYPES AND LUNG CANCER

	TCCA	CACA	CATA	CCCG	CCCA	CCTA	TACA	All others
Overall, *n* (case/control)	346/695	325/582	112/169	51/65	11/25	18/28	3/8	10/6
	1.0	1.2	1.4	1.8	0.81	1.6	1.3	3.2
Adj. OR[1]		(0.95, 1.5)	(1.0, 1.9)	(1.1, 2.8)	(0.35, 1.9)	(0.78, 3.1)	(0.31, 5.9)	(0.99, 11)
Stratified by smoking status in pack-years								
0	55/322	68/272	15/71	13/30	1/15	1/15	0/5	1/2
	1.0	1.2	0.98	1.6	0.31	0.46	–	3.1
Adj. OR[2]		(0.83, 1.9)	(0.50, 1.9)	(0.77, 3.5)	(0.039, 2.5)	(0.058, 3.7)		(0.26, 38)
>0–20	44/222	49/202	24/65	10/29	1/7	6/7	2/3	0/3
	1.0	1.3	1.8	1.7	0.55	4.5	3.8	–
Adj. OR[2]		(0.79, 2.0)	(1.0, 3.3)	(0.74, 3.8)	(0.064, 4.8)	(1.4, 15)	(0.59, 25)	
>20	247/151	208/108	73/33	28/6	9/3	11/6	1/0	9/1
	1.0	1.2	1.3	2.2	1.9	1.2	–	5.7
Adj. OR[2]		(0.85, 1.6)	(0.81, 2.1)	(0.85, 5.6)	(0.49, 7.1)	(0.44, 3.5)		(0.70, 46)

[1]Adjusted for gender, age, education, ethnicity and tobacco smoking (pack-years).–[2]Adjusted for gender, age, education and ethnicity.

TABLE VI – THE ASSOCIATIONS OF THE 4 *LIG1* HAPLOTYPES AND UADT CANCERS

	TCCA	CACA	CATA	CCCG	CCCA	CCTA	TACA	All others
Overall, *n* (case/control)	309/698	284/588	97/172	48/67	9/26	17/28	7/8	9/5
Adj. OR[1]	1.0	1.1	1.2	1.5	1.1	1.4	2.5	3.3
		(0.87, 1.3)	(0.92, 1.7)	(0.96, 2.3)	(0.49, 2.4)	(0.73, 2.8)	(0.86, 7.2)	(0.97, 11)
Stratified by smoking status in pack-years								
0	108/326	75/275	33/72	10/31	3/16	6/15	2/5	3/2
Adj. OR[2]	1.0	0.77	1.4	0.86	0.65	1.4	1.3	5.2
		(0.55, 1.1)	(0.86, 2.2)	(0.40, 1.9)	(0.18, 2.3)	(0.51, 3.7)	(0.25, 6.9)	(0.85, 32)
>0–20	70/222	74/204	25/67	8/30	5/7	1/7	3/3	0/2
Adj. OR[2]	1.0	1.2	1.1	0.73	2.4	0.52	3.2	–
		(0.80, 1.7)	(0.65, 2.0)	(0.31, 1.7)	(0.67, 8.3)	(0.062, 4.4)	(0.61, 17)	
>20	131/150	135/109	39/33	30/6	1/3	10/6	2/0	6/1
Adj. OR[2]	1.0	1.4	1.3	6.3	0.40	2.0	–	5.3
		(0.99, 2.1)	(0.74, 2.3)	(2.4, 17)	(0.039, 4.1)	(0.67, 6.1)		(0.58, 48)

[1]Adjusted for gender, age, education, ethnicity and tobacco smoking (pack-years).–[2]Adjusted for gender, age, education and ethnicity.

years old; OR = 3.1, 95% CL = 2.3, 4.1 for >16-years old) and UADT cancers (OR = 1.9, 95% CL= 1.5, 2.5 for ≤16-years old; OR = 1.3, 95% CL = 0.96, 1.6 for >16-years old) (Table II).

We did not detect departure from Hardy-Weinberg equilibrium for the genotyping frequencies among the controls. Adjusting for age, sex, race/ethnicity, education and pack-years of tobacco smoking, there was little or no association between the *LIG1* polymorphisms on exon 25 and exon 7 and lung-cancer incidence (Table III). Weak-to-moderate associations were observed between the *LIG1* SNPs on exon 2 and intron 1 and lung-cancer incidence, but these associations were not estimated precisely. The adjusted OR was 1.3 (95% CL = 0.97, 1.8) for the CT *vs.* CC genotype and 2.1 (95% CL = 0.74, 5.9) for the TT *vs.* CC genotype on exon 2 and 5.3 (95% CL = 0.92, 30) for the GG *vs.* AA genotype on intron 1. For the associations between the SNPs on exon 25, on exon 2 and on intron 1 and UADT cancer incidence (Table IV), the adjusted OR was 1.5 (95% CL = 1.0, 2.3) for the CC *vs.* TT genotype on exon 25, 1.3 (95% CL = 1.0, 1.8) for the CT *vs.* CC genotype and 2.0 (95% CL = 0.69, 5.8) for the TT *vs.* CC genotype on exon 2, and 5.9 (95% CL= 1.1, 31) for the GG *vs.* AA genotype on intron 1.

When stratifying on pack-years of smoking, we observed higher OR estimates among ever smokers (both ≤ 20 and >20 pack-years) than among never smokers for the associations of the SNPs on exons 25 and 7 with UADT cancers (Table IV), but this pattern was not observed in the other analyses (Tables III and IV). Furthermore, likelihood-ratio tests for the interactions on the multiplicative scale between the SNPs and pack-years of smoking yielded *p* values of 0.46 for exon 25 (rs20581), 0.99 for exon 7 (rs20580), 0.58 for exon 2 (rs20579) and 0.84 for intron 1 (rs439132) for lung cancer, and 0.31 for exon 25, 0.75 for exon 7, 0.82 for exon 2 and 0.29 for intron 1 for UADT cancers.

The haplotypes of the 4 *LIG1* polymorphisms were reconstructed, and the frequencies of the more common haplotypes (5% or higher) are shown in Tables V and VI. For lung cancer, when compared to the most common haplotype (rs20581-rs20580-rs20579-rs439132 = T-C-C-A), the adjusted OR was 1.2 (95% CL = 0.95, 1.5) for the CACA haplotype, 1.4 (1.0, 1.9) for the CATA haplotype and 1.8 (1.1, 2.8) for the CCCG haplotype. In addition, the association with the CATA haplotype was more evident among moderate smokers (>0–20 pack-years). The test for interaction between haplotypes and pack-years yielded a *p* value of 0.99.

For UADT cancers, when compared to the most common haplotype (rs20581-rs20580-rs20579-rs439132 = T-C-C-A), the adjusted OR was 1.1 (95% CL = 0.87, 1.3) for the CACA haplotype, 1.2 (0.92, 1.7) for the CATA haplotype and 1.5 (0.96, 2.3) for the CCCG haplotype. In addition, the associations with the CACA haplotype and the CCCG haplotype were more evident among heavy smokers (>20 pack-years) with the adjusted ORs of 1.4 (0.99, 2.1) and 6.3 (2.4, 17), respectively (Table VI). The test for interaction yielded a *p* value of 0.08.

Discussion

DNA ligase I (*LIG1*), located on chromosome 19,[13] is a 102-kDa nuclear enzyme that appears to be involved in joining Okazaki fragments during DNA replication and in sealing single-strand breaks in the base excision repair pathway.[13–15] Bentley *et al*. suggested that *LIG1* null cells have normal DNA repair capacity, but also have increased genome instability,[28] possibly affecting cancer susceptibility. It is important to identify the functional significance of the *LIG1* polymorphism to interpret its potential effects on cancer development in the population.[29,30] *LIG1* was thought to be essential for cell viability; however, Bentley *et al*. observed that cells lacking DNA ligase 1 are viable.[28] Bentley *et al*. hypothesized that the function of *LIG1* may be compensated by another ligase gene.[28] This hypothesis warrants further investigation.

DNA repair plays an essential role by protecting cells' chromosomes from environmental exposure, such as tobacco smoke.[31] Among the different repair pathways, base excision repair removes small lesions, DNA single strand breaks and nonbulky adducts.[31] An observed association between increased risk of UADT cancer and higher levels of BPDE-induced DNA adducts in peripheral lymphocytes may reflect a combination of carcinogen metabolism and DNA repair.[32] An increasing number of publications imply the association between reduced DNA repair capacity and altered cancer risk and the association between DNA repair genetic variations and DNA repair function and the consequently modified cancer risk.[33–36]

Only weak associations with the haplotypes of *LIG1* for lung cancer and associations with the genotypes of *LIG1* polymorphisms and the haplotypes for UADT cancer were observed in this study. For a few of the associations, the data showed more pronounced patterns among moderate and/or heavy smokers, consistent with our expectation to see stronger genotype associations among moderate/heavy smokers. In addition, interactions between smoking and the polymorphisms on the development of the above-mentioned cancers were implied, based on the stratified analyses. However, we had limited power to detect the extent of the potential interactions.

A previous study on the *LIG1* polymorphism on exon 7 (rs20580) and lung cancer detected no association between the *LIG1* polymorphism and the risk of lung cancer in a case-control study that included non-Hispanic whites (OR = 1.0, 95% CL 0.8, 1.5 for the AC *vs*. CC genotype; OR = 0.93, 95% CL 0.6, 1.4 for the AA *vs*. CC genotype) after adjusting for age, sex, smoking, alcohol use and family history of cancer.[20] In our study, we also investigated ethnic groups other than Caucasians. However, we did not have enough subjects to detect interactions with ethnicity, upon examining the combined associations of these 4 *LIG1* polymorphisms and tobacco smoking with lung and UADT cancers.

Nonetheless, when we considered all 4 of the SNPs together as haplotypes, we observed associations between the *LIG1* haplotypes and the cancer risks. Our findings showed the possibility that these 4 *LIG1* polymorphisms and haplotypes may play a role in the susceptibility to the cancers of interest. Tobacco smoking results in the formation of adducts on DNA.[37] The main pathway for repairing this type of DNA damage is NER.[30,38] Since *LIG1* is involved in the NER and BER pathways,[30] there is potential for interaction between this *SNP* and tobacco smoking. The associations of the *LIG1* polymorphisms and haplotypes with lung and head and neck cancers among moderate/heavy smokers may suggest carcinogenesis mechanisms involving exposure to tobacco smoke.

Another possible explanation for the observed association is that the examined SNPs could be in linkage disequilibrium with other functional SNPs that were not genotyped. Functional studies on the *LIG1* gene are warranted to test these hypotheses. Given the limited precision for interaction assessment in single studies, further studies of these effects are needed and study samples will have to be carefully pooled to obtain an accurate picture of gene-environment and gene–gene interactions. Furthermore, although diplotypes represent the chromosome pairs for a particular individual, and thus are more informative to assessing individual risk, we do not have enough power to assess diplotypes due to the limited sample size. Studies with larger sample sizes are warranted for further investigation on diplotypes.

A major strength of this study was that the study design was population-based. The controls were selected from the same population that gave rise to the cases based on a neighborhood-matching algorithm; therefore, it reduced an important potential source of selection bias.

Despite its strengths, this study also has a number of limitations. There might be possible selection bias introduced by nonparticipation of eligible persons, due to premature death, refusal or inability to contact. Lung cancer cases with poor prognoses may result in a rapid progression to an advanced stage, precluding interviews and participation in the study and resulting in possible selection bias. Nonetheless, the active surveillance procedure, which was used to identify cases shortly after diagnosis, reduces this problem. Recall bias may occur due to the nature in which information is collected in a case-control study. Compared to controls, cases might differentially recall exposures of tobacco smoking or alcohol consumption because of their cancer diagnosis. Reporting bias might occur if certain exposures are under or overreported, especially if behaviors or habits, such as tobacco smoking, carry a social stigma in the population. On the other hand, we expect minimal misclassification of the DNA repair genetic polymorphisms. Standard laboratory protocols were used to genotype the markers, while the laboratory personnel were blinded to the case-control status of the samples.

Additionally, our sample size, although modest, provided limited power to detect associations in subgroups, especially for the rare variants. We also assessed the false positive reporting probabilities (FPRPs) according to Wacholder *et al*.[39] With a prior of 25% for UADT cancer, the FPRP is 27% for Rs20581 (Exon 25) C/C *vs*. T/T; 24% for Rs20581 (Exon 25) C/C *vs*. T/T among heavy smokers; and 18% for Rs439132 (Intron 1) G/G *vs*. A/A. Thus, the probability of false positives among the reported associations is high.

In summary, we observed weak-to-moderate associations of the *LIG1* polymorphisms (on exon 2 and on intron 1) with lung cancer and of the *LIG1* polymorphisms (on exon 25, on exon 2 and on intron 1) with UADT cancers. Despite limitations of possible bias and imprecise estimation for rare genotypes and haplotypes, our findings suggest that *LIG1* polymorphisms and haplotypes may affect the risk of lung and UADT cancers. Although observing such effects may help us better understand the mechanisms through which smoking increases cancer risk, further research is needed to examine gene-smoking and gene–gene interactions in large populations.

Acknowledgements

The authors thank Mr. John Dermand for supervising the administration of questionnaires and the collection of buccal smears. The authors thank Ms. Taylor McAfee for extracting DNA from buccal cells, and Dr. Yan Cui for managing the data files. The authors are also indebted to the study participants for their dedication and commitment.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol 2001;2:533–43.
3. Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, Thun MJ. Cancer statistics, 2006. CA Cancer J Clin 2006;56:106–30.
4. IARC. Tobacco smoking. Lyon: IARC, 1986.
5. Ries LAG, Harkins D, Krapcho M, Mariotto A, Miller BA, Feuer EJ, Clegg L, Eisner MP, Horner MJ, Howlader N, Hayat M, Hankey BF, et al. SEER Cancer Statistics Review, 1975–2003 (based on November 2005 SEER data submission). 2006. Available at http://seer.cancer.gov/csr/1975_2003/ (accessed May 10, 2006).
6. Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg RS, Preston-Martin S, Bernstein L, Schoenberg JB, Stemhagen A, Fraumeni JF, Jr. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res 1988;48:3282–7.
7. Blot WJ, McLaughlin JK, Devesa SS. Cancers of the oral cavity and pharynx. In: Schottenfeld D, Fraumeni JF, eds. Cancer epidemiology and prevention. New York: Oxford University Press, 1996;666–80.
8. Austin DF, Reynolds P. laryngeal cancer. In: Schottenfeld D, Fraumeni JF, eds. Cancer epidemiology and prevention. New York: Oxford University Press, 1996;619–36.
9. Demokan S, Demir D, Suoglu Y, Kiyak E, Akar U, Dalay N. Polymorphisms of the XRCC1 DNA repair gene in head and neck cancer. Pathol Oncol Res 2005;11:22–5.
10. Huang WY, Olshan AF, Schwartz SM, Berndt SI, Chen C, Llaca V, Chanock SJ, Fraumeni JF, Jr, Hayes RB. Selected genetic polymorphisms in MGMT, XRCC1, XPD, and XRCC3 and risk of head and

neck cancer: a pooled analysis. Cancer Epidemiol Biomarkers Prev 2005;14:1747–53.
11. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. Science 2001;291:1284–9.
12. Berwick M, Vineis P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. J Natl Cancer Inst 2000; 92:874–97.
13. Lindahl T, Barnes DE. Mammalian DNA ligases. Annu Rev Biochem 1992;61:251–81.
14. Noguiez P, Barnes DE, Mohrenweiser HW, Lindahl T. Structure of the human DNA ligase I gene. Nucleic Acids Res 1992;20:3845–50.
15. Petrini JH, Xiao Y, Weaver DT. DNA ligase I mediates essential functions in mammalian cells. Mol Cell Biol 1995;15:4303–8.
16. Nocentini S. Rejoining kinetics of DNA single- and double-strand breaks in normal and DNA ligase-deficient cells after exposure to ultraviolet C and gamma radiation: an evaluation of ligating activities involved in different DNA repair processes. Radiat Res 1999;151: 423–32.
17. Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. Curr Biol 2000;10:919–22.
18. van Hoffen A, Balajee AS, van Zeeland AA, Mullenders LH. Nucleotide excision repair and its interplay with transcription. Toxicology 2003;193:79–90.
19. SNP500Cancer team members. SNP500Cancer Database. Cancer Genome Anatomy Project 2004 (cited 2004 Mar). Available at http://snp500cancer.nci.nih.gov/home.cfm.
20. Shen H, Spitz MR, Qiao Y, Zheng Y, Hong WK, Wei Q. Polymorphism of DNA ligase I and risk of lung cancer—a case-control analysis. Lung Cancer 2002;36:243–7.
21. Hashibe M, Morgenstern H, Cui Y, Tashkin DP, Zhang ZF, Cozen W, Mack TM, Greenland S. Marijuana use and the risk of lung and upper aerodigestive tract cancers: results of a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2006;15:1829–34.
22. Cui Y, Morgenstern H, Greenland S, Tashkin DP, Mao J, Cao W, Cozen W, Mack TM, Zhang ZF. Polymorphism of Xeroderma Pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. Int J Cancer 2006;118:714–20.
23. Cao W, Hashibe M, Rao JY, Morgenstern H, Zhang ZF. Comparison of methods for DNA extraction from paraffin-embedded tissues and buccal cells. Cancer Detect Prev 2003;27:397–404.
24. Tobler AR, Short S, Andersen MR, Paner TM, Briggs JC, Lambert SM, Wu PP, Wang Y, Spoonde AY, Koehler RT, Peyret N, Chen C, et al. The SNPlex genotyping system: a flexible and scalable platform for SNP genotyping. J Biomol Tech 2005;16:398–406.
25. Greenland S. Introduction to regression modeling. In: Rothman K, Greenland S, eds. Modern epidemiology. Philadelphia: Lippincott Williams & Wilkins, 1998;395–6.
26. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001; 68:978–89.
27. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162–9.
28. Bentley DJ, Harrison C, Ketchen AM, Redhead NJ, Samuel K, Waterfall M, Ansell JD, Melton DW. DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability. J Cell Sci 2002;115:1551–61.
29. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1513–30.
30. Mohrenweiser HW, Wilson DM, III, Jones IM. Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes. Mutat Res 2003;526:93–125.
31. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. Nature 2001;411:366–74.
32. Li D, Firozi PF, Chang P, Wang LE, Xiong P, Sturgis EM, Eicher SA, Spitz MR, Hong WK, Wei Q. In vitro BPDE-induced DNA adducts in peripheral lymphocytes as a risk factor for squamous cell carcinoma of the head and neck. Int J Cancer 2001;93:436–40.
33. Hall J, Hashibe M, Boffetta P, Gaborieau V, Moullan N, Chabrier A, Zaridze D, Shangina O, Szeszenia-Dabrowska N, Mates D, Janout V, Fabianova E, et al. The association of sequence variants in DNA repair and cell cycle genes with cancers of the upper aerodigestive tract. Carcinogenesis 2007;28:665–71.
34. 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.
35. Hu JJ, Mohrenweiser HW, Bell DA, Leadon SA, Miller MS. Symposium overview: genetic polymorphisms in DNA repair and cancer risk. Toxicol Appl Pharmacol 2002;185:64–73.
36. Spitz MR, Fueger JJ, Beddingfield NA, Annegers JF, Hsu TC, Newell GR, Schantz SP. Chromosome sensitivity to bleomycin-induced mutagenesis, an independent risk factor for upper aerodigestive tract cancers. Cancer Res 1989;49:4626–8.
37. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733–44.
38. Sancar A. DNA excision repair. Annu Rev Biochem 1996;65:43–81.
39. 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.