Cancer Research

Genetic Associations of 115 Polymorphisms with Cancers of the Upper Aerodigestive Tract across 10 European Countries: The ARCAGE Project

Cristina Canova, Mia Hashibe, Lorenzo Simonato, et al.

Cancer Res 2009;69:2956-2965. Published online April 1, 2009.

DOI:10.1158/0008-5472.CAN-08-2604

Research Article

Genetic Associations of 115 Polymorphisms with Cancers of the Upper Aerodigestive Tract across 10 European Countries: The ARCAGE Project

Cristina Canova,[1] Mia Hashibe,[2] Lorenzo Simonato,[1] Mari Nelis,[3] Andres Metspalu,[3,4] Pagona Lagiou,[5,6] Dimitrios Trichopoulos,[5,6] Wolfgang Ahrens,[7] Iris Pigeot,[7] Franco Merletti,[8] Lorenzo Richiardi,[8] Renato Talamini,[9] Luigi Barzan,[10] Gary J. Macfarlane,[11] Tatiana V. Macfarlane,[11] Ivana Holcátová,[12] Vladimir Bencko,[12] Simone Benhamou,[13,14] Christine Bouchardy,[15] Kristina Kjaerheim,[16] Ray Lowry,[17] Antonio Agudo,[18] Xavier Castellsagué,[18] David I. Conway,[19,20] Patricia A. McKinney,[20,21] Ariana Znaor,[22] Bernard E. McCartan,[23] Claire M. Healy,[23] Manuela Marron,[2] and Paul Brennan[2]

[1]Department of Environmental Medicine and Public Health, University of Padova, Padova, Italy; [2]IARC, Lyon, France; [3]University of Tartu, Institute of Molecular and Cell Biology/Estonian Biocentre; [4]The Estonian Genome Project of the University of Tartu, Tartu, Estonia; [5]University of Athens School of Medicine, Athens, Greece; [6]Harvard School of Public Health, Boston, Massachusetts; [7]Bremen Institute for Prevention Research and Social Medicine, University of Bremen, Bremen, Germany; [8]Unit of Cancer Epidemiology, Center for Experimental Research and Medical Studies and University of Turin, Turin, Italy; [9]Aviano Cancer Centre, Aviano, Italy; [10]General Hospital of Pordenone, Pordenone, Italy; [11]University of Aberdeen School of Medicine, Aberdeen, United Kingdom; [12]Charles University in Prague, 1st Faculty of Medicine, Institute of Hygiene and Epidemiology, Prague, Czech Republic; [13]Institut National de la Santé et de la Recherche Médicale, U794, Paris, France; [14]Centre National de la Recherche Scientifique, FRE2939, Institut Gustave Roussy, Villejuif, France; [15]Cancer Registry, Geneva, Switzerland; [16]Cancer Registry of Norway, Oslo, Norway; [17]University of Newcastle Dental School, Newcastle, United Kingdom; [18]Institut Català d'Oncologia, Barcelona, Spain; [19]University of Glasgow Dental School, Glasgow, Scotland; [20]Information Services Division, National Health Service, National Services Scotland, Edinburgh, Scotland; [21]University of Leeds, Leeds, United Kingdom; [22]Croatian National Cancer Registry, Croatian National Institute of Public Health, Zagreb, Croatia; and [23]Trinity College School of Dental Science, Dublin, Ireland

Abstract

Cancers of the upper aerodigestive tract (UADT) include malignant tumors of the oral cavity, pharynx, larynx, and esophagus and account for 6.4% of all new cancers in Europe. In the context of a multicenter case-control study conducted in 14 centers within 10 European countries and comprising 1,511 cases and 1,457 controls (ARCAGE study), 115 single nucleotide polymorphisms (SNP) from 62 *a priori*–selected genes were studied in relation to UADT cancer. We found 11 SNPs that were statistically associated with UADT cancers overall (5.75 expected). Considering the possibility of false-positive results, we focused on SNPs in *CYP2A6, MDM2, tumor necrosis factor* (*TNF*), and *gene amplified in squamous cell carcinoma 1* (*GASC1*), for which low *P* values for trend (*P* trend < 0.01) were observed in the main effects analyses of UADT cancer overall or by subsite. The rare variant of *CYP2A6* −47A>C (rs28399433), a phase I metabolism gene, was associated with reduced UADT cancer risk (*P* trend = 0.01). Three SNPs in the *MDM2* gene, involved in cell cycle control, were associated with UADT cancer. *MDM2* IVS5+1285A>G (rs3730536) showed a strong codominant effect (*P* trend = 0.007). The rare variants of two SNPs in the *TNF* gene were associated with a decreased risk; for *TNF* IVS1+123G>A (rs1800610), the *P* trend was 0.007. Variants in two SNPs of *GASC1* were found to be strongly associated with increased UADT cancer risk (for both, *P* trend = 0.008). This study is the largest genetic epidemiologic study on UADT cancers in Europe. Our analysis points to potentially relevant genes in various pathways. [Cancer Res 2009;69(7):2956–65]

Requests for reprints: Paul Brennan, International Agency for Research on Cancer, 150 cours Albert Thomas, 69008 Lyon, France. Phone: 33-4-72-73-8391; Fax: 33-4-72-73-83-42; E-mail: brennan@iarc.fr.

doi:10.1158/0008-5472.CAN-08-2604

Introduction

Cancers of the upper aerodigestive tract (UADT) include malignant tumors originating from the oral cavity, pharynx, larynx, and esophagus (1). Altogether, UADT cancers account for 5.2% of all new cancer cases worldwide and 6.4% in Europe (2, 3). Each year, there are more than 180,000 new cases and 105,000 deaths from cancer of the UADT in Europe; 32% of new cases occur in the oral cavity, 19% in the pharynx, 24% in the esophagus, and 25% in the larynx (2). Within Europe, incidence rates vary considerably, with high rates being observed in France, Italy, Spain, and countries in Central Europe, particularly in Hungary (2). This is thought to be mainly because of different patterns of alcohol and tobacco consumption and types of alcoholic beverage between European populations.

Established etiologic factors include tobacco consumption (both smoking and chewing) and heavy alcohol drinking (4, 5). Tobacco and alcohol jointly account for 80% of UADT cancer (6–8), but only a fraction of subjects exposed to these carcinogens will develop neoplastic lesions. Alcohol- and tobacco-related cancers are of special interest in molecular epidemiology because these exposures are thought to interact with a variety of genetic factors (9). There are gene polymorphisms involved in the metabolism of carcinogens from tobacco and alcohol in DNA repair and cell cycle control that may contribute to interindividual variation of risk. The study of single nucleotide polymorphisms (SNP) in UADT cancer etiology may help to identify high-risk subgroups and to better understand the pathways leading to these cancers. A crucial point is the appropriate selection of candidate genes and functionally relevant polymorphisms.

Alcohol is metabolized to acetaldehyde by alcohol dehydrogenases (ADH), whereas subsequent conversion of acetaldehyde to acetic acid is catalyzed by aldehyde dehydrogenases (ALDH). The major part of alcohol and aldehyde metabolism is carried out in the liver, but metabolism by ADH, the cytochrome *P*450–related

DOI:10.1158/0008-5472.CAN-08-2604

enzyme CYP2E1, and ALDH also occurs in the UADT. There is evidence that polymorphisms in the *ADH1B, ADH1C, ADH7*, and *ALDH2* genes are associated with UADT cancers (10–12).

Aromatic and heterocyclic amines, substances present in tobacco, require metabolic activation to interact with DNA. These major groups are metabolized by cytochrome *P*450–related enzymes, with *CYP1A1, CYP1A2, CYP2A6*, and *CYP2D6* being some of the most studied polymorphic genes for these enzymes. Other polymorphisms related to tobacco metabolism may also be relevant to UADT cancer risk. After *N*-acetylation, the *N*-hydroxyaromatic and *N*-hydroxyheterocyclic amines are further activated by *N*-acetyltransferase to acetoxy intermediates, which react with DNA to form adducts. Two *N*-acetyltransferase genes, *NAT1* and *NAT2*, are highly polymorphic. The *NAT2* gene is involved in the development of bladder and colorectal cancer (13). Another important group of phase II enzymes is the glutathione *S*-transferase family of genes, including *GSTM, GSTT*, and *GSTP* genes. For many of these genes, several studies on UADT cancers have been conducted in different populations, but results are very inconsistent (14–18).

Other potentially important sources of interindividual variability in relation to the development of cancer are DNA repair capacity and cell cycle control. The association between different variants in these pathways and UADT cancer has been investigated in multiple studies, but the results are inconsistent (19–25).

Here, we present a large study of >3,000 UADT cancer cases and controls recruited in 10 European countries that was conducted to investigate the association of genetic polymorphisms with the risk of these tumors. One of the major objectives of the study was to investigate the role of genetic variation with regard to the metabolism of alcohol and carcinogens from tobacco smoke, as well as DNA repair in the development of UADT cancers. We present the main effects of more than a hundred *a priori*–selected variants involved in various pathways, as well as some SNPs for which we have no prior knowledge of function. In addition, relevant stratified analyses investigating effect modification are presented.

Materials and Methods

Study population. ARCAGE (alcohol-related cancers and genetic susceptibility in Europe) is a multicenter case-control study conducted in 14 centers within 10 European countries: Prague (Czech Republic); Bremen (Germany); Athens (Greece); Aviano, Padova, and Turin (Italy); Dublin (Ireland); Oslo (Norway); Glasgow, Manchester, and Newcastle (United Kingdom); Barcelona (Spain); Zagreb (Croatia); and Paris (France). The IARC (Lyon, France) was responsible for the overall coordination of the study. The study has been described in more detail previously (26).

Briefly, according to the same protocol, each center recruited a group of newly diagnosed cases of UADT cancers (within 1 year of diagnosis), including oral cavity, pharynx, larynx, and esophagus, and a comparable group of hospital- or population-based controls. The subjects were recruited between January 2002 and December 2005, except for the French center where subjects were enrolled between 1987 and 1992. All cases were histologically or cytologically confirmed. Controls were frequency matched by sex, age, center, ethnicity, and referral (or residence) area to the case group. Hospital controls were selected from a strictly defined list of admission diagnoses by diseases unrelated to alcohol, tobacco, or dietary practices. The proportion of controls within a specific diagnostic group did not exceed 33%. In United Kingdom centers, control subjects were selected by random sampling from the registered list of persons in the general practice with which the case was registered.

The study was approved by the IARC ethics committee as well as the local ethics committees, and each participant provided written informed consent. Non-Caucasians were excluded from analysis because there were only 13 of them and adjustment for ethnicity would have resulted in sparse cells. Both cases and controls underwent a face-to-face interview on current and previous alcohol consumption, dietary habits, tobacco consumption, and other lifestyle factors using the same questionnaire. Overall, the average participation rate was 82% among cases and 68% among controls. A total of 2,304 cases and 2,227 controls were included in ARCAGE, but DNA for genotyping was available for 1,788 cases and 1,685 controls. Of the remaining 1,058 subjects, 801 did not accept to give blood and DNA quality was not good enough for genotyping for 257.

Selection of relevant polymorphisms. *A priori*, we decided to include any SNP that satisfied at least one of the following criteria: (*a*) allele frequency >5% in Caucasians, (*b*) a previously reported association with lung or UADT cancer ($P < 0.05$), or (*c*) low or unknown allele frequency in Caucasians but high allele frequency in non-Caucasians (>25%). We selected SNPs following the third criteria due to the incompleteness of the SNP databases at the beginning of the study. We reasoned that if a disease-related SNP is common in non-Caucasian population, but information on its prevalence in the Caucasian population was limited, it was still of high enough priority to be included. An initial search was conducted at the start of the study in 2002 to identify all genes with some preliminary data of an association with UADT cancers or genes that may have an association based on their supposed biological function. The review was conducted using overviews of genetic susceptibility and cancer (27, 28) and gene expression analysis of UADT cancer (29). In addition, a Medline search using keywords such as "genetic," "susceptibility," and "cancer" was performed. This review resulted in a list of 92 genes, which were considered to be the most likely candidates in tobacco- or alcohol-related cancer susceptibility.

Before genotyping, a review of the literature about polymorphisms of the relevant genes was carried out. The list of polymorphisms was expanded with information from more recent literature and publicly available databases (National Center for Biotechnology Information dbSNP,[24] HGVbase,[25] and SNP Consortium[26]). Information recorded for each polymorphism included its exact position with respect to a reference Genbank sequence, its nature (SNP, insertion/deletion, repeat), its known functional effect (e.g., amino acid change), the allelic frequencies in various human ethnic groups, and the flanking sequence. The final result was 554 SNPs from 92 genes all relevant to UADT cancers and meeting our inclusion criteria.

In total, two to three common SNPs were selected per gene. SNPs in the coding regions of the genes were preferred to the noncoding SNPs, as well SNPs with some published evidence of functionality (30, 31).

These selection criteria resulted in a final group of 166 SNPs from 76 genes to be included on the ARCAGE-ARRAY. This was a compromise between the suitable SNPs, candidate genes, and the possible volume of the genotyping.

Genotyping. Genotyping of all subjects, using an array-based methodology, was conducted by the arrayed primer extension (APEX) method (32, 33). Microarray spotting and genotyping by APEX reactions were conducted as described (32). Signal detection and interpretation were performed by Genorama imaging and analysis software (Asper Biotech Ltd.[27]).

Validation of ARCAGE-ARRAY. For the array validation, 125 (some of them in duplicate) anonymous human DNA samples from the SNP500 collection (so-called Coriell samples) were genotyped for the selected 166 polymorphisms. The samples included subjects from the main ethnic groups represented in the United States, with a subset of Centre d'Etude du Polymorphisme Humain families. The genotype data of SNP500 collection samples generated by bidirectional sequencing are publicly available on the Web site.[28] Markers that did not satisfy quality control criteria were replaced with neighboring markers.

[24] http://www.ncbi.nlm.nih.gov/projects/SNP
[25] http://www.hgvbase.org/
[26] http://snp.cshl.org
[27] http://www.asperbio.com
[28] http://snp500cancer.nci.nih.gov

DOI:10.1158/0008-5472.CAN-08-2604

Table 1. Selected characteristics of 1,511 cases and 1,457 controls from 14 European centers in the ARCAGE study

	Cases		Controls	
	n	%	*n*	%
Center				
Prague	142	9.40	113	7.76
Bremen	161	10.66	167	11.46
Athens	183	12.11	148	10.16
Aviano	136	9.00	139	9.54
Padova	112	7.41	103	7.07
Turin	137	9.07	152	10.43
Dublin	27	1.79	12	0.82
Oslo	111	7.35	114	7.82
Glasgow	41	2.71	47	3.23
Manchester	123	8.14	149	10.23
Newcastle	80	5.29	92	6.31
Barcelona	77	5.10	78	5.35
Zagreb	46	3.04	36	2.47
Paris	135	8.93	107	7.34
Sex				
Male	1,229	81.34	1,104	75.77
Female	282	18.66	353	24.23
Age (y)				
<45	102	6.75	1,436	9.81
45–49	159	10.52	127	8.72
50–54	208	13.77	219	15.03
55–59	311	20.58	254	17.43
60–64	251	16.61	217	14.89
65–69	209	13.83	212	14.55
70–74	163	10.79	170	11.67
75+	108	7.15	115	7.89
Smoking status				
Never	163	10.79	484	33.22
Former	368	24.35	501	34.39
Current	980	64.86	472	32.40
Alcohol intake (average drinks/day)				
Never	87	5.76	162	11.12
<1 drink per day	365	24.16	561	38.50
1–2 drinks per day	422	27.93	474	32.53
3–4 drinks per day	280	18.53	167	11.46
5+ drinks per day	357	23.63	93	6.38
Subsite				
Oral cavity/oropharynx	733	48.51		
Larynx/hypopharynx	591	39.11		
Esophagus	185	12.24		
Overlapping	2	0.13		
Total	1,511		1,457	

Data analysis. There were 1,788 cases and 1,685 controls with DNA available. Subjects with a call rate below 90% for the 166 SNPs (that is <150 SNPs called) were excluded. This led to the inclusion of 3,044 subjects (cases and controls) in the analysis. Some 10% of the studied subjects were randomly selected and their DNA samples were reanalyzed for each polymorphism to evaluate the concordance of the genotyping. SNPs with less than 95% concordance and 95% call rates were excluded from the analysis. This led to the exclusion of 39 SNPs.

Departure from Hardy-Weinberg equilibrium (HWE) in the controls was tested by a χ^2 test using a type I error probability of $\alpha = 0.01$. We excluded SNPs with $P < 0.01$ (rather than $P < 0.05$) because of the anticonservatism of this test, as noted by Wigginton and colleagues (34). This resulted in the exclusion of a further 12 SNPs, leaving 115 SNPs for data analysis.

Demographic variables and factors associated with UADT cancer, including center, sex, age at diagnosis (in 5-y groups), smoking, and alcohol habits, were included in the analysis. With respect to smoking, subjects were classified as never, former, or current smokers. Cumulative tobacco consumption was calculated as pack-years (i.e., the product of smoking duration in years and intensity in packs per day). With respect to alcohol drinking, subjects were classified as never drinkers and drinkers of <1, 1 to 2, 3 to 4, and ≥5 drink equivalents per day. The definition for one drink equivalent was 14 g ethanol, which approximately corresponds to 150 mL wine, 330 mL beer, and 36 mL spirits (5). Drinks per day were calculated by summing each type of alcohol in drink-years and dividing this by the total duration of alcohol drinking. Seventy-six subjects with missing information on at least one of the matching and adjustment variables were excluded from the analyses [i.e., age (7 missing), sex (1 missing), smoking habit

DOI:10.1158/0008-5472.CAN-08-2604

(pack-years, 14 missing), and drinking habit (drink per day, 65 missing)]. Analyses eventually included 1,511 cases and 1,457 controls.

Unconditional logistic regression was used to calculate odds ratios (OR) and 95% confidence intervals (95% CI) adjusted by country, sex, and age at diagnosis (in 5-y groups) as categorical variables and cumulative tobacco consumption (pack-years) and alcohol intake (daily number of drinks) as continuous variables. Both genotype and allele ORs were calculated. Allele ORs are depicted in Fig. 2. Trend test for ordered variables, notably genotypes, were conducted by treating heterozygotes as being at intermediate risk. A Wald test was performed to test the statistical significance of each coefficient (β) in the model [compared with a χ^2 with 1 degree of freedom (*df*).

The quantile-quantile (Q-Q) plot was constructed by ranking the observed χ^2 statistics with 1 *df* for the main analysis from smallest to largest and plotting them against the expected values $[-\log(i/(L+1))]$, where L is the number of SNPs. Deviations from the line of equality ($y = x$) correspond to SNPs that deviate from the null hypothesis of no association at any locus.

Analyses were stratified by country, smoking status (never, former, and current), alcohol intake (none or light drinkers of <1 drink per day, moderate drinkers of 1 to <3 drinks per day, and heavy drinkers of 3+ drinks per day), and subsite (oral cavity/oropharynx, larynx/hypopharynx, and esophagus).

Forest plots were used to present the stratified results (allele ORs) for the SNPs with a *P* value for trend of <0.01 overall or by subsite.

Heterogeneity among the stratified ORs was assessed by conducting a likelihood ratio test comparing a model that included the product terms between the genetic variant and the stratification variable and a model without a product term. As ORs for different subsites were not mutually independent because the controls overlap, the *Q* test for heterogeneity was calculated instead of χ^2.

All tests were two sided, and a *P* value of <0.05 was considered to be statistically significant. The statistical analyses were conducted with Statistical Analysis System software (version 9.1; SAS Institute). Linkage disequilibrium between variants was tested in the controls by measure of R^2 using STATA software (version 8). The false-positive report probability (FPRP) for statistically significant observations was estimated using the procedure recommended by Wacholder and colleagues (35).

Results

Table 1 shows the demographic characteristics of the 1,511 cases and 1,457 controls included in the analyses. The age distribution was comparable between cases and controls, whereas more males were recruited among cases. The average number of subjects recruited by each center was 108 for the cases and 104 for the controls. As expected, the proportion of current smokers was higher among cases than among controls, as well as the average number of drinks per day (3.2 drinks in cases; 1.7 drinks in controls).

Figure 1 presents the comparison between the observed and expected distribution of the trend test statistic (χ^2) adjusted for possible confounders. We observed 11 SNPs with *P* values for trend test (*P* trend) of <0.05 against 5.75 expected under the null hypothesis of no association at any locus. By subsite, 10 SNPs with statistically significant *P* trend were found for esophageal cancer, 8 for oral cavity/oropharynx, and 4 for larynx/hypopharynx cancer.

Results on the main effects for the 115 SNPs are presented in Tables 2 to 4, grouped by presumed gene function.

Among phase I metabolizing genes (Table 2), we found two SNPs associated with the risk for UADT cancer. The rare variant allele of *CYP2A6* −47A>C (rs28399433) was associated with a reduced risk for UADT cancer (*P* trend = 0.01), whereas the rare variant allele of *CYP2C8* intron 9 G>A (rs1934951) was found to be associated with an increased risk of UADT (*P* trend = 0.02). We did not find any significant associations among SNPs in phase II metabolizing genes and genes coding for ADH and ALDH enzymes shown in Table 2.

Among the DNA repair genes (Table 3), two SNPs involved in nucleotide excision repair were associated with a reduced UADT cancer risk: *ERCC1* IVS5+33A>C (rs3212961) for the homozygous variant genotype (OR, 0.45; 95% CI, 0.23–0.90) and *ERCC4* S835S (rs1799801) for the heterozygous genotype (OR, 0.83; 95% CI, 0.70–0.98; Table 3). For two SNPs involved in the base excision repair gene, *XRCC1* P206P (rs915927) and *XRCC3* IVS5-571A>G (rs861530), the corresponding rare alleles were associated with reduced UADT cancer risk (*P* trend = 0.04 for both).

We found four significant associations between SNPs involved in cell cycle control and UADT cancers risk (Table 4). The *CDKN1A* A>G (rs2395655) variant was associated with an increased risk of UADT particularly for the homozygous variant (OR, 1.26; 95% CI, 0.99–1.59; *P* trend = 0.045). Three SNPs in the *MDM2* gene were associated with UADT cancer risk. In particular, the rare allele of *MDM2* IVS5+1285A>G (rs3730536) showed a strong codominant effect (*P* trend = 0.007). The three *MDM2* variants were in linkage disequilibrium among the controls (R^2 range: 0.36–0.95).

As shown in Table 4, the rare alleles of two SNPs in the tumor necrosis factor (*TNF*) gene, particularly the *TNF* IVS1+123G>A (rs1800610) variant (*P* trend = 0.007), were associated with a decreased UADT cancer risk. These two SNPs were in weak linkage equilibrium (R^2 = 0.05).

The two variants of gene amplified in squamous cell carcinoma 1 (*GASC1*) gene were found to be strongly associated with increased risk of UADT cancer (*P* trend = 0.008 for both the SNPs; R^2 = 0.97; Table 4). Additionally, the *LZTS1* G>A (rs3735836) heterozygous variant was associated with increased UADT cancer risk (OR, 1.40; 95% CI, 0.99–1.97; *P* trend = 0.036); two cases but no controls were homozygous for the rare variant. With respect to esophageal cancer (data not shown in Table 4), variant alleles of *deleted in esophageal cancer 1* (*DEC1*) and *deleted in lung and esophageal cancer 1* (*DLEC1*) genes were associated with a decreased risk of this cancer [*DEC1* 5′UTR G>A (rs3750505): heterozygotes OR, 0.50; 95% CI, 0.29–0.85; *P* trend = 0.06; *DLEC1* UTR C>T (rs2073401): *P* trend = 0.03].

Figure 2 shows stratified results for the four SNPs for which we found strong significant associations with UADT cancer risk (*P* trend < 0.01, overall or by subsite): *CYP2A6* −47A>C, *MDM2* IVS5+1285A>G, *TNF* IVS1+123G>A, and *GASC1* intron 10 G>T. Only

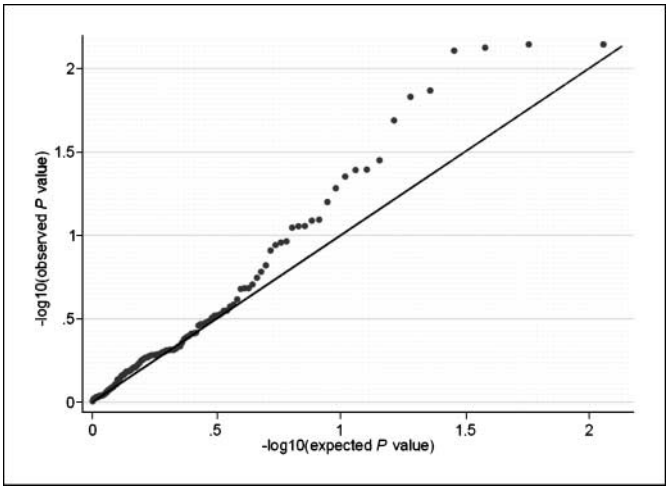

Figure 1. Q-Q plot for the trend statistics (Cochran-Armitage 1 *df* χ^2 trend test), adjusted for sex, age in quinquennia, and country as categorical variables and pack-year and alcohol intake as continuous variables.

DOI:10.1158/0008-5472.CAN-08-2604

Table 2. Main effects of SNPs in carcinogen metabolism genes for upper aerodigestive cancer

Function	SNP name rs number	Homozygotes common allele		Heterozygotes			Homozygotes rarer allele			*P* trend
		Ca	Co	Ca	Co	OR (95% CI)	Ca	Co	OR (95% CI)	
ADH and ALDH	ADH1B Ex9+77A>G rs17033	1,267	1,239	212	191	1.06 (0.84–1.33)	8	7	1.35 (0.46–3.94)	0.530
	ADH1C IVS6-892A>G rs1662058	486	497	734	676	1.06 (0.88–1.27)	258	239	1.12 (0.88–1.42)	0.331
	ADH1C T151Trs2241894	854	805	517	508	0.93 (0.79–1.11)	85	86	1.01 (0.72–1.42)	0.625
	ALDH2 Ex1+82A>G rs886205	951	951	470	419	1.04 (0.87–1.24)	58	52	1.09 (0.72–1.67)	0.563
	ALDH2 intron 8 G>A rs4646777	992	980	454	416	1.00 (0.84–1.19)	60	53	1.08 (0.71–1.63)	0.830
	ALDH2 intron 8 G>C rs4646776	1,505	1,456	4	—	—	1	—	—	0.960
Phase I	CYP1A1 I462F rs1048943	1,388	1,325	114	119	0.86 (0.64–1.15)	3	6	0.44 (0.10–1.99)	0.182
	CYP1A1 IVS1+606T>G rs2606345	636	606	667	664	0.99 (0.83–1.17)	175	158	1.03 (0.80–1.35)	0.907
	CYP1A2 IVS1-154A>C rs762551	702	682	638	603	1.03 (0.87–1.22)	140	152	0.83 (0.63–1.09)	0.445
	CYP1A2 N516N rs2470890*	594	583	638	610	1.04 (0.88–1.24)	223	210	0.97 (0.76–1.24)	1.000
	CYP2A6 −47A>C rs28399433	1,346	1,245	157	193	**0.76 (0.59–0.97)**	6	11	0.55 (0.19–1.56)	**0.014**
	CYP2C8 intron 9 G>A rs1934951	966	998	421	364	1.16 (0.97–1.39)	64	46	1.51 (0.99–2.30)	**0.021**
	CYP2C9 I359L rs1057910	1,297	1,254	198	183	1.07 (0.84–1.35)	7	10	0.62 (0.21–1.78)	0.888
	CYP3A4 IVS10+12G>A rs2242480	1,215	1,189	276	248	1.06 (0.87–1.31)	11	15	0.62 (0.26–1.48)	0.919
	EPHX1 H139R rs2234922	1,000	992	421	399	1.00 (0.84–1.19)	80	54	1.39 (0.95–2.05)	0.287
	MPO I717V rs2759	1,430	1,379	77	65	1.03 (0.71–1.48)	1	1	0.24 (0.01–4.52)	0.944
Phase II	COMT L136L rs4818	522	477	700	734	0.86 (0.72–1.03)	264	232	1.13 (0.90–1.43)	0.655
	COMT V158M rs4680	367	356	737	732	0.90 (0.74–1.09)	393	357	0.99 (0.80–1.24)	0.967
	GSTA2 exon 2 G>C rs1803684	1,310	1,273	191	165	1.12 (0.88–1.42)	7	7	1.05 (0.35–3.22)	0.394
	GSTA2 UTR A>G rs2254050	662	612	633	627	0.91 (0.77–1.08)	164	187	0.79 (0.61–1.02)	0.064
	GSTM3 Ex8+91G>A rs7483	752	697	607	627	0.89 (0.75–1.05)	136	123	1.03 (0.77–1.38)	0.527
	GSTP1 A114V rs1138272	1,298	1,249	193	189	1.03 (0.81–1.30)	10	11	0.87 (0.35–2.16)	0.945
	GSTP1 I105V rs1695	647	649	673	600	1.16 (0.98–1.37)	151	156	0.98 (0.75–1.28)	0.471
	GSTT2 −21225G>C rs140190	483	469	679	691	0.93 (0.78–1.12)	309	261	1.13 (0.91–1.42)	0.405
	NAT2 G286E rs1799931	1,424	1,382	77	71	1.14 (0.80–1.64)	1	1	0.45 (0.01–15.53)	0.525
	NAT2 R268K rs1208	467	442	742	736	0.93 (0.78–1.11)	286	259	1.01 (0.80–1.28)	0.943
	NQO1 exon 5 CT rs1131341	1,366	1,319	118	122	0.94 (0.71–1.25)	1	5	0.12 (0.01–1.14)	0.300
	SOD2 −1221G>A rs2758346	382	396	774	700	1.17 (0.97–1.42)	336	339	0.98 (0.78–1.22)	0.925
	SULT1A1 exon 7 AG rs1801029	718	708	645	592	1.10 (0.93–1.30)	129	138	0.93 (0.70–1.23)	0.824

NOTE: OR adjusted for sex, age in quinquennia, and country as categorical variables and pack-year and alcohol intake as continuous variables. In bold, statistically significant results, $P < 0.05$.

Abbreviations: Ca, cases; Co, control.

*P value for HWE, 0.01–0.05.

one SNP is presented for the *GASC1* gene because the two variants were in strong linkage disequilibrium.

As shown in Fig. 2, for *CYP2A6* −47A>C (rs28399433), we observed a strong association of the variant allele with esophageal cancer (OR, 0.32; 95% CI, 0.16–0.64; $P = 0.001$) and less strong associations with cancer of the oral cavity/oropharynx (OR, 0.76; 95% CI, 0.58–1.01; $P = 0.057$) and larynx/hypopharynx (OR, 0.88; 95% CI, 0.65–1.19; $P = 0.41$), with a significant test for heterogeneity between the ORs by subsites ($Q = 7.01$; $P = 0.03$). The inverse association of the A allele with UADT cancer was highly significant among heavy drinkers (OR, 0.47; 95% CI, 0.31–0.71), with a significant test for heterogeneity ($\chi^2 = 8.91$; $P = 0.01$). We did not observe significant heterogeneity in the stratified results with respect to the *MDM2* variant.

We observed a considerably stronger association of *TNF* IVS1+123G>A (rs1800610; OR, 0.67; 95% CI, 0.50–0.89) with UADT cancer among heavy drinkers than among moderate drinkers, but the χ^2 for heterogeneity was not significant ($P = 0.11$). We found a stronger positive association of the variant rare allele of *GASC1* intron 10 G>T (rs818912) with the risk of esophageal cancer (OR, 1.42; 95% CI, 1.07–1.87; $P = 0.011$) than with risk of oral cavity/oropharyngeal cancer (OR, 1.21; 95% CI, 1.02–1.41; $P = 0.04$) or laryngeal/hypopharyngeal cancer (OR, 1.07; 95% CI, 0.89–1.28; $P = 0.53$).

The associations of *CYP2A6* −47A>C, *MDM2* IVS5+1285A>G, *TNF* IVS1+123G>A, and *GASC1* intron 10 G>T with UADT cancer were likely to be true positives (FPRP ≤ 0.2) when the prior probability was at least 10% (FPRP: 0.125, 0.060, 0.061, and 0.067, respectively) but not with a prior probability of 1% (FPRP: 0.611, 0.412, 0.419, and 0.441, respectively; ref. 35).

Discussion

We studied 115 SNPs involved in various pathways in relation to UADT cancer in 1,511 cases and 1,457 controls recruited in 10 European countries.

Eleven SNPs were associated with UADT cancers, whereas 5 or 6 were expected. More statistically significant associations were

DOI:10.1158/0008-5472.CAN-08-2604

found for esophageal cancer risk than for oral cavity/oropharynx and larynx/hypopharynx, but the possibility of false-positive findings because of the small number of esophageal cases should be taken into account. Considering the possibility of false-positive results, our stratified analysis focused on four SNPs in *CYP2A6, MDM2, TNF*, and *GASC1*, for which low *P* values for trend (*P* trend < 0.01) were observed in the main effects analyses overall or by subsites. The *CYP2A6* gene is responsible for the metabolic activation of *N*-nitrosamines (36). A recent review by Kamataki and colleagues (37) suggested that genetic polymorphisms of *CYP2A6* may affect cancer risk in smokers but not in nonsmokers. In our data, we found evidence of a stronger effect in heavy drinkers ($P = 0.01$). No interaction by smoking status was evident. Tan and colleagues (38) showed that the *CYP2A6* gene deletion is associated with an increased risk of lung and esophageal cancer but not with a reduced tendency to smoke. The CYP2A6*4 deletion is, however, rare in European populations, with a prevalence of between 0% and 5% (39, 40), and will have a minimal effect on the association with other more common CYP2A6 variants. The apparent inverse association of *CYP2A6* −47A>C was also more evident for esophageal cancer compared with the other sites. A recent study (41) on lung cancer found a similar inverse association of the rare variant of the −47A>C polymorphism, irrespective of smoking status. Even in the absence of information on the CYP2A6 gene deletion, our results still support a role for this SNP.

MDM2 is a target gene of the transcription factor tumor protein p53. Overexpression of this gene can inactivate the protein p53, diminishing its tumor suppressor function. We have found different *MDM2* polymorphisms to be associated with either increased or decreased UADT cancer risk. The rs1695146 and rs2701092 SNPs are in strong linkage disequilibrium ($R^2 = 0.95$) but less so with rs3730536 SNP, the rare variant of which showed the strongest inverse association. No association with UADT cancer risk was found for *P53* gene.

To our knowledge, there has been only one published report on *MDM2* polymorphisms and UADT cancer risk (25). The authors reported that the rare variants of both the *P53* 72Arg>Pro and the *MDM2* 309T>G polymorphisms were associated with an increased risk of esophageal cancer with an apparent multiplicative interaction between the two SNPs. Pine and colleagues (42) reported that the rare variant of the *MDM2* SNP at position 309 was not strongly associated with lung cancer risk, whereas Li and colleagues (43) found that the same variant was associated with decreased risk of lung cancer.

We found apparent inverse associations for two polymorphisms in the *TNF* gene, a proinflammatory cytokine with both procarcinogenic and anticarcinogenic proprieties (44). A recent study (45) on the association between TNF and breast cancer risk concluded that the *TNF* IVS1+123G>A rs1800610 SNP was not significantly associated with breast cancer. To the best of our knowledge, there have been no published reports on *TNF* IVS1+123G>A rs1800610, *TNF* −1210C>T rs1799964, and UADT cancer risk.

Two variants of the *GASC1* gene were found to be strongly associated with increased risk of UADT cancer, particularly

Table 3. Main effects of SNPs in DNA repair genes for upper aerodigestive cancer

SNP name rs number	Homozygotes common allele		Heterozygotes			Homozygotes rarer allele			*P* trend
	Ca	Co	Ca	Co	OR (95% CI)	Ca	Co	OR (95% CI)	
APEX intron 4 T>C rs3136817	826	803	537	515	1.00 (0.84–1.19)	92	94	1.02 (0.74–1.42)	0.927
ERCC1 IVS3+74C>G rs3212948	626	613	666	634	1.07 (0.90–1.27)	190	192	1.02 (0.79–1.31)	0.665
ERCC1 IVS5+33A>C rs3212961	1,091	1,061	390	356	1.07 (0.90–1.29)	17	29	**0.45 (0.23–0.90)**	0.729
ERCC1 N118N rs11615 rs3177700	538	532	699	632	1.14 (0.95–1.35)	215	215	1.02 (0.80–1.30)	0.554
ERCC4 R415Q rs1800067	1,253	1,228	236	209	1.07 (0.86–1.34)	13	10	1.13 (0.46–2.78)	0.496
ERCC4 S835S rs1799801	767	697	576	603	**0.83 (0.70–0.98)**	143	122	1.06 (0.80–1.41)	0.343
LIG3 Ex21-250C>T rs1052536	449	429	742	728	0.97 (0.81–1.16)	291	281	0.92 (0.73–1.16)	0.491
LIG3 IVS9-761A>G rs3135998	515	505	733	682	1.07 (0.89–1.27)	239	238	1.02 (0.81–1.30)	0.708
MGMT IVS2+21006C>T rs913118*	491	448	652	659	0.89 (0.74–1.06)	321	309	0.90 (0.72–1.12)	0.286
MGMT I143V rs2308321	1,164	1,098	303	320	0.83 (0.68–1.00)	33	25	1.19 (0.67–2.10)	0.210
MGMT L53L rs1803965*	1,136	1,105	337	304	1.01 (0.84–1.23)	28	34	0.74 (0.42–1.28)	0.646
OGG1 IVS3-845A>G rs3219008*	926	949	495	426	1.19 (1.00–1.41)	70	69	1.07 (0.73–1.55)	0.116
PCNA IVS2-124C>T rs25406	499	469	746	710	1.02 (0.86–1.22)	247	259	0.95 (0.75–1.20)	0.740
POLB intron 3 C>A rs3136718	1,304	1,243	176	189	0.91 (0.72–1.16)	8	11	0.70 (0.25–1.92)	0.336
POLB IVS1-89C>T rs3136717	1,182	1,146	308	279	1.07 (0.88–1.31)	10	18	0.56 (0.24–1.32)	0.928
XRCC1 intron 5 G>A rs762507	513	450	686	688	0.88 (0.73–1.05)	248	259	0.82 (0.65–1.04)	0.082
XRCC1 P206P rs915927	498	424	732	743	**0.83 (0.69–0.99)**	265	269	0.81 (0.64–1.02)	**0.041**
XRCC1 Q399R rs25487	636	600	652	672	0.94 (0.80–1.12)	190	152	1.24 (0.96–1.60)	0.351
XRCC2 IVS2-3377G>A rs3218518	468	449	698	685	1.00 (0.83–1.20)	284	282	0.98 (0.78–1.23)	0.850
XRCC3 IVS5-571A>G rs861530	696	628	638	627	0.89 (0.75–1.05)	138	157	0.77 (0.59–1.01)	**0.041**
XRCC9 intron 13 C>T rs587118	373	363	714	721	0.97 (0.80–1.18)	401	356	1.15 (0.92–1.43)	0.212
XRCC9 intron 2 T>G rs504082	753	728	608	588	0.98 (0.83–1.16)	132	122	0.95 (0.71–1.27)	0.743

NOTE: OR adjusted for sex, age in quinquennia, and country as categorical variables and pack-year and alcohol intake as continuous variables. In bold, statistically significant results, $P < 0.05$.
*P value for HWE, 0.01–0.05.

DOI:10.1158/0008-5472.CAN-08-2604

Table 4. Main effects of SNPs in genes from other pathways for upper aerodigestive cancer

Function	SNP name rs number	Homozygotes common allele		Heterozygotes			Homozygotes rarer allele			*P* trend
		Ca	Co	Ca	Co	OR (95% CI)	Ca	Co	OR (95% CI)	
Addiction	ANKK1 E713K rs1800497	1,020	968	411	424	0.91 (0.76–1.08)	55	51	1.12 (0.73–1.71)	0.610
	DRD2 Ex8+246C>T rs6276*	742	698	621	630	1.00 (0.85–1.18)	111	105	1.06 (0.78–1.46)	0.789
	DRD2 P319P rs6277	482	429	725	752	0.86 (0.72–1.03)	299	268	1.05 (0.83–1.32)	0.962
Adhesion	CDH1 A692A rs1801552	587	565	687	683	0.99 (0.84–1.18)	204	174	1.22 (0.95–1.57)	0.245
	CDH11 type2 intron 1 G>A rs35178	652	608	661	673	0.92 (0.78–1.09)	168	163	0.88 (0.68–1.15)	0.263
	CDH2 intron 8 T>C rs597591	408	382	761	735	0.97 (0.80–1.17)	338	333	0.93 (0.74–1.16)	0.505
	CDH2 intron 8 T>C rs665781	424	408	749	729	0.99 (0.82–1.19)	312	306	0.98 (0.78–1.23)	0.855
	CTNNB1 D780D rs2293303	1,497	1,439	6	4	1.03 (0.27–3.87)	1	—	—	0.630
	ITGA3 intron 23 G>A rs756847	1,212	1,175	277	265	0.98 (0.80–1.21)	14	10	1.87 (0.78–4.46)	0.679
Apoptosis	TNF IVS1+123G>A rs1800610*	1,094	1,017	331	351	**0.79 (0.65–0.96)**	42	46	0.69 (0.43–1.11)	**0.007**
	TNF −1210C>T rs1799964	976	884	440	481	0.84 (0.71–1.00)	71	73	0.89 (0.62–1.29)	0.089
	TNFAIP2 intron 5 A>G rs2234131	610	541	682	695	0.89 (0.75–1.06)	196	182	0.98 (0.76–1.26)	0.518
	TNF-α -487A>G rs3091256 = rs1800629*	1,081	1,068	386	345	1.12 (0.93–1.35)	41	42	0.91 (0.57–1.47)	0.483
	TNF-α exon 4 A>C rs3093665	1,406	1,357	96	90	1.03 (0.74–1.42)	2	—	—	0.659
Cell communication	GJB5 UTR A>G rs2275229	602	557	655	639	0.91 (0.77–1.09)	210	217	0.85 (0.67–1.08)	0.153
Cell cycle control	CDKN1A A>G rs2395655	446	473	767	719	1.15 (0.96–1.38)	284	249	1.26 (0.99–1.59)	**0.045**
	CDKN1A p21 Ex3+70C>T rs1059234	1,281	1,222	209	217	0.85 (0.68–1.06)	12	11	1.01 (0.42–2.46)	0.210
	CDKN1A p21 S31R rs1801270	1,296	1,239	189	195	0.89 (0.70–1.12)	13	9	1.34 (0.53–3.39)	0.538
	CDKN1A p21 UTR T>A rs3176336	485	500	727	671	1.11 (0.93–1.33)	285	271	1.11 (0.88–1.39)	0.307
	CDKN2A Ex4+83C>T rs3088440	1,310	1,229	186	205	0.83 (0.66–1.05)	12	9	1.09 (0.42–2.83)	0.199
	MDM2 intron 4 G>A rs3730532 rs2701092	516	526	676	676	1.06 (0.88–1.26)	249	212	**1.39 (1.09–1.77)**	0.015
	MDM2 intron 8 G>A rs1695146	539	536	712	689	1.06 (0.89–1.26)	246	220	**1.29 (1.02–1.63)**	0.053
	MDM2 IVS5+1285A>G rs3730536	610	518	651	677	**0.77 (0.64–0.91)**	199	199	0.78 (0.61–1.00)	**0.007**
	TP53 P72R rs1042522	811	768	565	565	0.90 (0.76–1.06)	116	110	1.01 (0.75–1.36)	0.494
	WDR79 F150F rs2287498	1,280	1,212	204	220	0.84 (0.67–1.05)	12	8	1.39 (0.55–3.57)	0.270
Folate	MTHFR IVS10+35A>G rs1476413	822	784	580	560	1.03 (0.87–1.21)	101	107	0.83 (0.60–1.14)	0.567
	MTHFR A222V rs1801133	599	582	655	643	1.00 (0.85–1.19)	188	187	0.89 (0.69–1.15)	0.492
	MTHFR Q594R rs2274976	1,386	1,327	119	121	0.97 (0.73–1.30)	3	3	1.13 (0.21–6.03)	0.893
	TYMS E127D rs3786362	1,507	1,447	2	6	0.23 (0.04–1.26)	—	—	—	0.091
Signaling	LTBP1 Ala126Ala rs1065324	385	370	742	704	1.03 (0.85–1.24)	375	374	0.93 (0.75–1.16)	0.533
	TGFB1 308bp 3° of STP C>T rs1800469*	632	602	697	673	1.02 (0.86–1.21)	159	146	1.15 (0.87–1.52)	0.389
	TGFB1 intron T>C rs3826714	970	937	463	461	0.95 (0.80–1.13)	59	47	1.33 (0.86–2.06)	0.790
	TGFB1 IVS2+114C>T rs2241716	1,457	1,404	15	18	0.90 (0.43–1.89)	4	—	—	0.528
	TGFBR2 intron 6 G>A rs3773649	797	778	581	562	1.00 (0.85–1.19)	123	100	1.14 (0.84–1.56)	0.545
	TGFBR2 IVS1-5868A>G rs3863057	654	630	679	661	1.04 (0.88–1.23)	163	155	0.98 (0.75–1.28)	0.929
	TGFBR2 IVS2-1369A>G rs891595	1,117	1,105	331	304	1.04 (0.86–1.26)	24	23	1.05 (0.55–1.97)	0.696
	TGFBR2 IVS6+236C>T rs2276768	1,208	1,139	267	270	0.94 (0.76–1.15)	16	16	0.80 (0.38–1.70)	0.413
Tumor suppressor	FHIT G>A rs931322*	473	451	727	740	0.97 (0.81–1.16)	285	242	1.08 (0.86–1.37)	0.625
	FHIT intron 2 C>T rs994929	721	677	628	614	0.97 (0.82–1.14)	120	129	0.84 (0.62–1.13)	0.305
	FHIT intron 3 G>A rs610766	422	383	721	711	0.91 (0.75–1.10)	358	344	0.92 (0.74–1.15)	0.468
	FHIT intron 3 T>C rs1038516	986	930	453	467	0.95 (0.80–1.12)	58	53	1.01 (0.67–1.52)	0.665
Other	DEC1 5°UTR G>A rs3750505	1,220	1,141	251	258	0.94 (0.76–1.15)	23	21	1.20 (0.64–2.26)	0.833
	DEC1 intron 7 T>C rs2269701*	845	894	524	441	1.17 (0.98–1.39)	78	76	0.95 (0.66–1.36)	0.316
	DEC1 UTR T>C rs1537728	634	637	669	651	1.05 (0.88–1.24)	189	156	1.27 (0.98–1.65)	0.110
	DLEC1 intron 11 T>C rs2236630	621	567	621	642	0.85 (0.71–1.01)	190	196	0.90 (0.70–1.16)	0.167
	DLEC1 intron 9 A>G rs2070484	925	865	496	508	0.86 (0.72–1.02)	87	75	1.07 (0.75–1.51)	0.345
	DLEC1 UTR C>T rs2073401	1,338	1,249	161	200	**0.75 (0.59–0.96)**	9	4	3.01 (0.85–10.67)	0.125
	DLEC1 UTR G>A rs2073712	792	781	531	522	0.94 (0.79–1.12)	117	96	1.16 (0.85–1.59)	0.803
	ERBB2 IVS4-61C>T rs1810132	735	676	639	640	0.88 (0.74–1.04)	118	122	0.82 (0.61–1.12)	0.083
	ERBB2 P1170A rs1058808	665	650	685	640	0.98 (0.83–1.16)	153	156	0.90 (0.68–1.18)	0.492
	GASC1 5°UTR A>G rs913589	410	406	724	724	0.98 (0.81–1.18)	348	307	1.11 (0.89–1.39)	0.393
	GASC1 intron 10 G>T rs818912	866	925	558	455	**1.27 (1.07–1.50)**	76	69	1.25 (0.87–1.81)	**0.008**
	GASC1 N396D rs2296067*	863	921	558	448	**1.31 (1.10–1.55)**	80	76	1.19 (0.83–1.70)	**0.008**
	IL1R1 intron 11 A>G rs3917318*	744	706	620	618	1.01 (0.86–1.19)	106	96	1.08 (0.79–1.49)	0.699
	LZTS1 5°UTR G>C rs3735839*	679	674	675	603	1.04 (0.88–1.23)	157	177	0.93 (0.72–1.21)	0.867
	LZTS1 G>A rs3735836	1,407	1,376	90	72	1.40 (0.99–1.97)	2	—	—	**0.036**

(*Continued on the following page*)

DOI:10.1158/0008-5472.CAN-08-2604

Table 4. Main effects of SNPs in genes from other pathways for upper aerodigestive cancer (Cont'd)

Function	SNP name rs number	Homozygotes common allele		Heterozygotes			Homozygotes rarer allele			*P* trend
		Ca	Co	Ca	Co	OR (95% CI)	Ca	Co	OR (95% CI)	
	LZTS1 intron 3 C>T rs904000	518	512	711	688	1.03 (0.86–1.22)	259	240	1.07 (0.85–1.34)	0.586
	LZTS1 intron 3 T>G rs903998	1,014	973	443	427	0.97 (0.81–1.15)	47	51	0.90 (0.58–1.41)	0.596
	RAF1 intron 8 C>G rs1039244	886	886	532	477	1.17 (0.99–1.38)	76	73	1.15 (0.80–1.65)	0.089
	RARB intron 1 T>C rs755661	492	452	683	670	0.88 (0.74–1.06)	283	289	0.84 (0.67–1.06)	0.112
	RARB intron 4 G>A rs1286756	545	488	697	718	0.85 (0.72–1.02)	255	236	0.98 (0.77–1.24)	0.507
	RNF6 exon 6 T>C rs301047	1,016	960	396	379	0.96 (0.80–1.15)	47	48	0.88 (0.56–1.38)	0.525
	RNF6 intron 6 A>G rs301058	457	455	741	695	1.12 (0.94–1.34)	304	289	1.08 (0.86–1.35)	0.422
	SDC1 intron 2 G>A rs1106111	606	577	705	688	0.99 (0.84–1.17)	191	173	1.12 (0.86–1.44)	0.546

NOTE: OR adjusted for sex, age in quinquennia, and country as categorical variables and pack-year and alcohol intake as continuous variables. In bold, statistically significant results, $P < 0.05$.
*P value for HWE, 0.01–0.05.

esophageal cancer. *GASC1* is located at 9p24, a region frequently amplified in tumor tissues of esophageal squamous cell carcinomas (46). Yang and colleagues cloned *GASC1* and showed that *GASC1* was overexpressed in cell lines. Using several independent lines of evidence, Cloos and colleagues (47) have shown that *GASC1* and its homologues *JMJD2A* and *JMJD2B* demethylate the repressive histone H3K9me3/me2 marks both *in vitro* and *in vivo*. These findings indicate that histone trimethylation is a reversible modification and may potentially have far-reaching implications for human disease, particularly cancer. Only one association study has been published on *GASC1* variants and cancer risk, notably a genome-wide case-control study of esophageal cancer using GeneChip Mapping 10K array (48). In that study, the rs1340513 SNP in the *GASC1* gene was found to be associated with esophageal cancer risk. Based on HAPMAP data, there is a strong association between this variant and rs818912 in Europeans ($R^2 = 0.96$).

In addition to the four SNPs described above, there was evidence of an association ($0.01 < P < 0.05$ with main effects considering both a dominant and a codominant model) of *CYP2C8* intron 9 G>A rs1934951, *ERCC1* IVS5+33A>C rs3212961, *ERCC4* S835S rs1799801, *XRCC1* P206P rs915927, *XRCC3* IVS5-571A>G rs861530, *CDKN1A* A>G rs2395655, *DLEC1* UTR C>T rs2073401, and *LZTS1* G>A rs3735836 with UADT cancer risk. Replication of these associations in other large studies is warranted.

Three of nine ADH and ALDH SNPs were excluded from the analyses due to deviation from HWE or low call rates, particularly *ADH1B* H48R (rs1229984), for which a recent study of about 3,800 cases and 5,200 controls found results that were strongly significant (12).

Use of hospital controls, in all but the United Kingdom centers, would have been considered a limitation for traditional case-control studies. However, only controls with admission diagnoses unrelated to diet, alcohol, or smoking were eligible to participate. Furthermore, the main disadvantage of the case-control studies, that is, susceptibility to bias when estimating effects of exposures that are measured retrospectively, does not generally apply when studying genetic effects and statistical interaction between genotype and environmental exposures (49). In any case, the genetic associations were generally similar for the United Kingdom and the other centers, as well as allele frequencies. Population stratification is unlikely to have affected our results, given that we excluded from the analysis people who were not of European ethnicity, and allele frequencies did not differ substantially between countries. Genome-wide data from tobacco-related cancers within Europe that adjust for genetic markers of ethnicity show only limited evidence of genetic mixture within countries and little potential for population stratification (50). Genome-wide data for the ARCAGE study have recently become available as part of a large initiative on ~70% of participants using the Illumina 317K panel. Adjustment by genetic markers of ethnicity shows little or no change to results adjusted by country (data not shown).

Strengths of our study include the use of a standard protocol, the implementation of strict criteria for control selection, the use of a single laboratory for all DNA extraction, and the use of state-of-the-art procedures for genetic determination, as well as detailed assessment of lifetime tobacco and alcohol consumption. We also included a large number of cases and controls, allowing for more statistical power in main effect and stratified analyses (26).

This study is the largest genetic epidemiologic study on UADT cancers with biological samples available in Europe. Our analysis points to associations with genes in various pathways. Further detailed characterization of these pathways and replication in other large studies of head and neck cancer will be essential.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 7/30/08; revised 1/7/09; accepted 1/30/09.

Grant support: European Commission's 5th Framework Program (Contract No. QLK1-2001-00182), Italian Association for Cancer Research, Compagnia di San Paolo/FIRMS, Region Piemonte, and Padova University (Contract No. CPDA057222).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank all the patients and their families for their participation and the clinicians and staff of the hospitals, interviewers, data managers, pathology departments, and primary care clinics for their support.

DOI:10.1158/0008-5472.CAN-08-2604

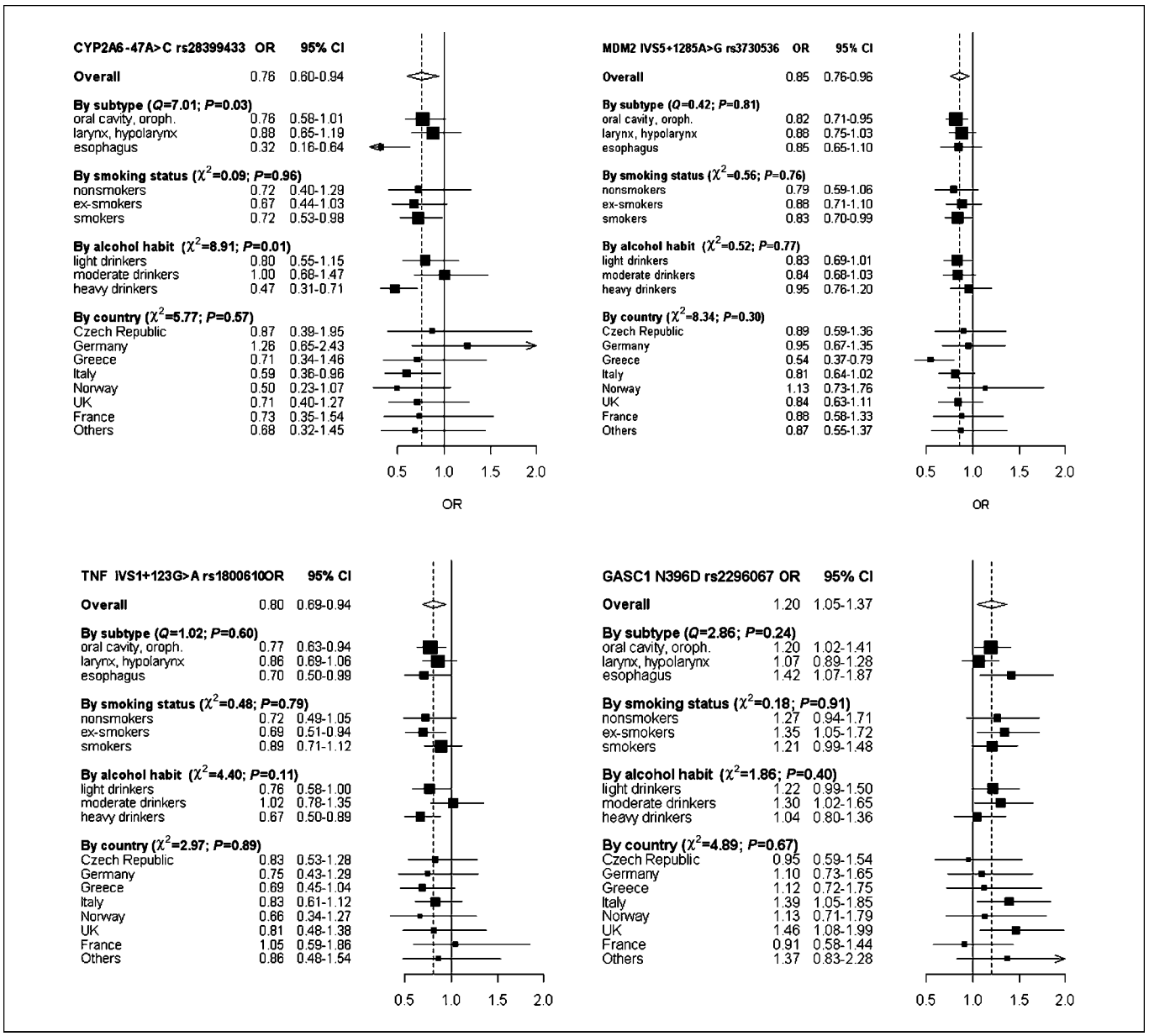

Figure 2. Forest plot of the *CYP2A6* −47A>C (rs28399433), *MDM2* IVS5+1285A>G (rs3730536), *TNF* IVS1+123G>A (rs1800610), and *GASC1* N396D (rs2296067) SNPs: OR and 95% CIs comparing rare with common allele; OR adjusted for sex, age in quinquennia, and country as categorical variables and pack-year and alcohol intake as continuous variables, where appropriate.

References

1. Dobrossy L. Epidemiology of head and neck cancer: magnitude of the problem. Cancer Metastasis Rev 2005; 24:9–17.
2. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: cancer incidence, mortality and prevalence worldwide, version 2.0. IARC CancerBase No. 5. Lyon (France): IARC Press; 2004.
3. Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J. Cancer incidence in five continents, vol. VII. IARC Scientific Publications No. 143. Lyon (France): IARC Press; 1997.
4. IARC monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 83, Lyon, Tobacco smoke and involuntary smoking, Lyon (France): IARC Press; 2004.
5. IARC monographs on the evaluation of carcinogenic risks to humans, vol. 96. Alcoholic beverage consumption and ethyl carbamate (urethane). Lyon (France): IARC Press; 2007.
6. Blot WJ, McLaughlin JK, Winn DM, et al. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res 1988;48:3282–7.
7. Franceschi S, Talamini R, Barra S, et al. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and oesophagus in northern Italy. Cancer Res 1990;50:6502–7.
8. Castellsague X, Quintana MJ, Martinez MC, et al. The role of type of tobacco and type of alcoholic beverage in oral carcinogenesis. Int J Cancer 2004;108:741–9.
9. Hung RJ, van der Hel O, Tavtigian SV, Brennan P, Boffetta P, Hashibe M. Perspectives on the molecular epidemiology of aerodigestive tract cancers. Mutat Res 2005;592:102–18.
10. Hashibe M, Boffetta P, Zaridze D, et al. Evidence for an important role of alcohol- and aldehyde-metabolizing genes in cancers of the upper aerodigestive tract. Cancer Epidemiol Biomarkers Prev 2006;15:696–703.
11. Peters ES, McClean MD, Liu M, Eisen EA, Mueller N, Kelsey KT. The ADH1C polymorphism modifies the risk of squamous cell carcinoma of the head and neck associated with alcohol and tobacco use. Cancer Epidemiol Biomarkers Prev 2005;14:476–82.
12. Hashibe M, McKay JD, Curado MP, et al. Multiple ADH genes are associated with upper aerodigestive cancers. Nat Genet 2008;40:707–9.
13. Sanderson S, Salanti G, Higgins J. Joint effects of the *N*-acetyltransferase 1 and 2 (NAT1 and NAT2) genes and smoking on bladder carcinogenesis: a literature-based systematic HuGE review and evidence synthesis. Am J Epidemiol 2007;166:741–51.

DOI:10.1158/0008-5472.CAN-08-2604

14. Morita S, Yano M, Tsujinaka T, et al. Association between genetic polymorphisms of glutathione *S*-transferase P1 and *N*-acetyltransferase 2 and susceptibility to squamous-cell carcinoma of the oesophagus. Int J Cancer 1998;79:517–20.
15. Olshan AF, Weisser MC, Watson MA, Bell DA. GSTM1, GSTT1, GSTP1, CYP1A1, and NAT1 polymorphisms, tobacco use, and the risk of head and neck cancer. Cancer Epidemiol Biomarkers Prev 2000;9:185–91.
16. Geisler SA, Olshan AF. GSTM1, GSTT1, and the risk of squamous cell carcinoma of the head and neck: a mini-HuGE review. Am J Epidemiol 2001;154: 95–105.
17. Hashibe M, Brennan P, Strange RC, et al. Meta- and pooled analyses of GSTM1, GSTT1, GSTP1, and CYP1A1 genotypes and risk of head and neck cancer. Cancer Epidemiol Biomarkers Prev 2003;12:1509–17.
18. Sugimura T, Kumimoto H, Tohnai I, 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: 11–8.
19. Sturgis EM, Castillo EJ, Li L, et al. Polymorphisms of DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis 1999;20:2125–9.
20. Olshan AF, Watson MA, Weissler MC, Bell DA. XRCC1 polymorphisms and head and neck cancer. Cancer Lett 2002;178:181–6.
21. Sturgis EM, Castillo EJ, Li L, et al. XPD/ERCC2 exon 8 polymorphisms: rarity and lack of significance in risk of squamous cell carcinoma of the head and neck. Oral Oncol 2002;38:475–7.
22. Shen H, Sturgis EM, Dahlstrom KR, Zheng Y, Spitz MR, Wei Q. A variant of the DNA repair gene XRCC3 and risk of squamous cell carcinoma of the head and neck: a case-control analysis. Int J Cancer 2002;99: 869–72.
23. Zheng Y, Shen H, Sturgis EM, et al. Haplotypes of two variants in p16 (CDKN2/MTS-1/INK4a) exon 3 and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Epidemiol Biomarkers Prev 2002;11:640–5.
24. Shen H, Zheng Y, Sturgis EM, Spitz MR, Wei Q. P53 codon 72 polymorphism and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Lett 2002;183:123–30.
25. Hong Y, Miao X, Zhang X, et al. The role of P53 and MDM2 polymorphisms in the risk of oesophageal squamous cell carcinoma. Cancer Res 2005;65: 9582–7.
26. Lagiou P, Georgila C, Minaki P, et al. Alcohol-related cancers and genetic susceptibility in Europe—the ARCAGE project: study samples and data collection. Eur J Cancer Prev 2009;18:76–84.
27. Gemignani F, Landi S, Vivant F, Zienolddiny S, Brennan P, Canzian F. A catalogue of polymorphisms related to xenobiotic metabolism and cancer susceptibility. Pharmacogenetics 2002;12:459–63.
28. Landi S, Gemignani F, Gioia-Patricola L, Chabrier A, Canzian F. Evaluation of a microarray for genotyping polymorphisms related to xenobiotic metabolism and DNA repair. Biotechniques 2003;35:816–20, 822, 824–7.
29. Al Moustafa AE, Alaoui-Jamali MA, Batist G, et al. Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. Oncogene 2002;21: 2634–40.
30. Syvänen AC. Accessing genetic variation: genotyping single nucleotide polymorphisms. Nat Rev Genet 2001;2: 930–42.
31. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234–6.
32. Kurg A, Tõnisson N, Tollett J, Georgiou I, Shumaker J, Metspalu A. Arrayed primer extension: solid phase four-color DNA resequencing and mutation detection technology. Genet Test 2000;4:1–7.
33. Tõnisson N, Kurg A, Kaasik K, Lõhmussaar E, Metspalu A. Unravelling genetic data by arrayed primer extension. Clin Chem Lab Med 2000;38:165–70.
34. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76:887–93.
35. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability of false positive reports in molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.
36. Guengerich FP, Shimada T. Oxidation of toxic and carcinogenic chemicals by human cytochrome P-450 enzymes. Chem Res Toxicol 1991;4:391–407.
37. Kamataki T, Fujieda M, Kiyotani K, Iwano S, Kunitoh H. Genetic polymorphism of CYP2A6 as one of the potential determinants of tobacco-related cancer risk. Biochem Biophys Res Commun 2005;338: 306–10.
38. Tan W, Chen GF, Xing DY, Song CY, Kadlubar FF, Lin DX. Frequency of CYP2A6 gene deletion and its relation to risk of lung and oesophageal cancer in the Chinese population. Int J Cancer 2001;95:96–101.
39. Nakajima M, Yoshida R, Fukami T, McLeod HL, Yokoi T. Novel human CYP2A6 alleles confound gene deletion analysis. FEBS Lett 2004;569:75–81.
40. Fukami T, Nakajima M, Sakai H, McLeod HL, Yokoi T. CYP2A7 polymorphic alleles confound the genotyping of CYP2A6*4A allele. Pharmacogenomics J 2006;6:401–12.
41. Gemignani F, Landi S, Szeszenia-Dabrowska N, et al. Development of lung cancer before the age of 50: the role of xenobiotic metabolizing genes. Carcinogenesis 2007;28:1287–93.
42. Pine SR, Mechanic LE, Bowman ED, et al. MDM2 SNP309 and SNP354 are not associated with lung cancer risk. Cancer Epidemiol Biomarkers Prev 2006;15:1559–61.
43. Li G, Zhai X, Zhang Z, Chamberlain RM, Spitz MR, Wei Q. MDM2 gene promoter polymorphisms and risk of lung cancer: a case-control analysis. Carcinogenesis 2006;27:2028–33.
44. Balkwill F. Tumor necrosis factor or tumor promoting factor? Cytokine Growth Factor Rev 2002;13:135–41.
45. Gaudet MM, Egan KM, Lissowska J, et al. Genetic variation in tumor necrosis factor and lymphotoxin-α (TNF-LTA) and breast cancer risk. Hum Genet 2007;121: 483–90.
46. Yang ZQ, Imoto I, Fukuda Y, et al. Identification of a novel gene, GASC1, within an amplicon at 9p23-24 frequently detected in oesophageal cancer cell lines. Cancer Res 2000;60:4735–9.
47. Cloos PA, Christensen J, Agger K, et al. The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature 2006;442:307–11.
48. Hu N, Wang C, Hu Y, et al. Genome-wide association study in oesophageal cancer using GeneChip mapping 10K array. Cancer Res 2005;65:2542–6.
49. Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. Lancet 2001;358:1356–60.
50. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008; 452:633–7.