Carcinogenesis vol.32 no.12 pp.1855–1861, 2011
doi:10.1093/carcin/bgr211
Advance Access publication September 16, 2011

Population-specific genetic associations with oesophageal squamous cell carcinoma in South Africa

Hannah Bye[1], Natalie J.Prescott[1], Marco Matejcic[2], Elizabeth Rose[3], Cathryn M.Lewis[1], M.Iqbal Parker[2] and Christopher G.Mathew[1,*]

[1]Department of Medical and Molecular Genetics, King's College London, King's Health Partners, Guy's Hospital, London SE1 9RT, UK, [2]International Centre for Genetic Engineering and Biotechnology and Division of Medical Biochemistry, University of Cape Town, Cape Town, South Africa and [3]Medical Research Council, Tygerberg, South Africa

*To whom correspondence should be addressed. Tel: +44 (0) 20 7188 3713; Fax: +44 (0) 207 188 2585;
Email: christopher.mathew@kcl.ac.uk

Genetic variants in multiple cellular pathways have been associated with an altered risk of oesophageal cancer. In this study, eight genes previously associated with an altered risk of oesophageal squamous cell carcinoma (OSCC) in European or Asian populations were investigated in two South African populations. We genotyped 12 single-nucleotide polymorphisms and one insertion/deletion variant in 1463 individuals from the Black and Mixed Ancestry populations. No polymorphisms were associated with OSCC in the Black population. In the Mixed Ancestry population, *ALDH2* +82 G > A (rs886205) was significantly associated with a reduced risk of OSCC (odds ratio = 0.70, 95% confidence interval = 0.55–0.89; $P = 0.0038$). Several other polymorphisms showed a suggestive association ($P < 0.05$), including *ADH1B* Arg48His (rs1229984), *COX-2* −1195G > A (rs689466), *CASP8* Asp302His (rs1045485) and *MGMT* Leu84Phe (rs12917). Haplotype analysis indicated that the *FAS* polymorphisms −670 A > G (rs1800682) and −1377 G > A (rs2234767) were both associated with OSCC in the Mixed Ancestry population ($P = 0.006$ and $P = 0.004$, respectively), as well as the *CASP8* (−652 6Ndel:302His) haplotype ($P = 0.0013$). This study indicates several instances of population-specific differences in the genetic etiology of OSCC between these two South African populations and between them and other high-risk populations, which may reflect differences in their ancestry and environmental exposures.

Introduction

Oesophageal cancer is the eighth most common cancer in the world and is responsible for >300 000 deaths a year (1). The disease has a very poor prognosis with a 5 years survival rate of <10% (2). Two main subtypes exist, squamous cell carcinoma and adenocarcinoma, which are etiologically unrelated. Oesophageal squamous cell carcinoma (OSCC) is the predominant form in developing countries (3). High-risk regions have been identified in China, Japan, Iran and southern Africa. In the Eastern Cape Province of South Africa, oesophageal cancer is the most common malignancy in Black males and the second most common in Black females, with an incidence of 32.7 and 20.2 cases per 100 000 people, respectively (4). Alcohol and tobacco are implicated in the majority of cases in the western world (1). In South Africa, additional risk factors include nutritional deficiencies, consumption of maize contaminated with the *Fusarium* fungus and human papilloma virus infection [reviewed in Hendricks *et al.* (2)].

Most genetic studies in OSCC have focused on candidate genes involved in alcohol metabolism, detoxification of carcinogens, DNA repair, apoptosis and cell proliferation [reviewed in Lao-Sirieix *et al.* (5)]. However, the results have not always been consistent, particularly across different populations. This may reflect differences in the prevalence of susceptibility variants between populations, differences in environmental exposures or technical issues such as small sample sizes which are not well powered to detect modest genetic effects. Genome-wide association studies (GWAS) in Japanese and Chinese populations have detected association of genetic variants in *ADH1B*, *ALDH2*, *PLCE1* and *C20orf54* with OSCC (6–8). Recently, a GWAS in upper aerodigestive cancers including OSCC in European populations reported associations in *ADH7*, the *ALDH2* locus and a novel association in the DNA repair gene *HEL308* (9). Our previous studies in the South African population have detected association of genetic variants in several genes with OSCC, including *GSTP1* (10), *CYP2E1* (11), *SULT1A1* and *CYP3A5* (12). In this study, we have sought to obtain a clearer understanding of genetic and environmental factors contributing to the pathogenesis of OSCC in an expanded cohort from the Black and Mixed Ancestry populations of South Africa by investigation of 12 single-nucleotide polymorphisms (SNPs) and one insertion/deletion variant from eight genes with previous robust evidence of association with OSCC in other populations.

Abbreviations: CI, confidence interval; OR, odds ratio; GWAS, genome-wide association study; LD, linkage disequilibrium; OR, odds ratio; OSCC, oesophageal squamous cell carcinoma; PCR, polymerase chain reaction.

Materials and methods

Study subjects

A total of 1463 individuals were recruited from the Black and Mixed Ancestry populations of South Africa. The Black subjects were mainly Xhosa-speakers from the Eastern or Western Cape of South Africa, who are one of the major populations originating from the Bantu-speaking peoples of Southern Africa. The Mixed Ancestry subjects were from the Western Cape. This population (also referred to in the literature and self-reported as the 'coloured' population of South Africa) is an admixed population with major ancestral components from the indigenous Khoisan, Bantu-speaking Africans, Europeans and Asians (13). The study consisted of 358 OSCC patients and 477 controls from the Black population and 201 OSCC patients and 427 controls from the Mixed Ancestry population. All patients were recruited between March 2000 and September 2010 at Groote Schuur Hospital (GSH), Cape Town, South Africa, with histologically confirmed primary invasive OSCC. Control samples were recruited from the same populations as the patients and from the same geographical area, age group, gender and ethnic group. Data on alcohol and tobacco use were available for OSCC cases. Smoking status was subdivided into those who were current smokers, former smokers or never-smokers. Drinkers were defined as subjects who consumed alcohol at least once in every week. Demographic and exposure data are given in Table I. Whole blood samples were collected with informed consent from all subjects and DNA was extracted at the University of Cape Town. Ethical approval for the study was obtained from the joint University of Cape Town/GSH Research Ethics Committee.

Candidate genes

Polymorphisms were selected for genotyping following a literature review of published genetic association studies involving OSCC and other head-and-neck squamous cell carcinomas (Table II). A total of 12 SNPs and one insertion/deletion variant were genotyped in cases and controls: *ADH1B* Arg48His (rs1229984); *ADH7* Gly92Ala (rs1573496); *ALDH2* Glu504Lys (rs671), +82 A > G (rs886205) and −261 C > T (rs441); *FAS* −670 G > A (rs1800682) and −1377 G > A (rs2234767); *FASL* −844 T > C (rs763110); *COX-2* −765 G > C (rs20417) and −1195 A > G (rs689466); *MGMT* Leu84Phe (rs12917); *CASP8* Asp302His (rs1045485) and *CASP8* −652 6N ins/del (rs3834129). It is important to note that the nomenclature for these variants is based on older versions of the human genome and nucleotide annotation. However, for simplicity and to allow consistency with previous literature in the field, we have maintained this 'common nomenclature' throughout. A summary of the updated nomenclature for these variants using the more recent genome annotation (NCBI36) adopted by the Human Genome Variation Society (www.HGVS.org) is provided in Supplementary Table I, available at *Carcinogenesis* Online.

SNP genotyping

All samples were genotyped using the TaqMan 5′ exonuclease assay with primers and probes designed and synthesized by Applied Biosystems (Carlsbad, CA) (19).

SNPs in the alcohol and aldehyde dehydrogenase genes (with the exception of *ALDH2* −261 T > C) were genotyped using validated TaqMan drug metabolism genotyping assays (Applied Biosystems). SNPs in *MGMT*, *FAS* and *COX-2* −1195 A > G were genotyped using validated TaqMan SNP-genotyping assays (Applied Biosystems). Custom TaqMan assays were designed for *COX-2* −765G > C [forward primer, CCCCCTCCTTGTTTCTTGGAA; reverse primer, TGCTTAG GACCAGTATTATGAGGAGAA; reporter ACCTTTCCC(G/C)CCTCTC], *CASP8* Asp302His [forward primer, ACCACGACCTTTGAAGAGCTT; reverse primer, TCCATGAGTTGGTAGATTTTCAAAATCTCA; reporter CCCCAC(G/C)]ATGACTG] and *ALDH2* −261 G > C (forward primer, AGCCTGGGTGCCAGAGAGA; reverse primer, CCTGACAGCATTCACTTAGAACAAC; reporter 1, CTCGGCCTCAAAA; reporter 2, ACTCGGTCTCAAAAA). Reactions were carried out in 2.5 µl volumes in 96-well plates. Each reaction contained 20 ng DNA, Absolute QPCR ROX mix (Abgene, Epsom, UK) and SNP assay mix (Applied Biosystems) according to assay instructions and were performed on a PTC-0225 DNA Engine (MJ Research, Waltham, MA). Fluorescent levels at the polymerase chain reaction (PCR) end-point were determined using a 7900HT Fast Real-Time PCR system (Applied Biosystems) and genotypes assigned using SDS 2.2.2 software (Applied Biosystems).

Insertion/deletion genotyping

Primers for the *CASP8* −652 6N ins/del were previously designed by Sun *et al.* (25). Briefly, the 5 µl PCR reaction contained 1× PCR mastermix (Promega, Madison, WI), 0.4 µM of each primer (Sigma, Dorset, UK), 10 ng of DNA and was performed on a thermocycler as above. The PCR products were separated by capillary electrophoresis on an ABI3730xl DNA Analyzer (Applied Biosystems) and sized using GeneMapper software (Applied Biosystems).

Statistical analysis

Pearson's chi-squared (χ^2) test was used to determine deviations from the Hardy–Weinberg equilibrium; all genotype frequencies were in Hardy–Weinberg equilibrium in both populations, with the exception of *FAS* -670 G > A in OSCC cases from the Black population ($P = 0.027$). Genotype and allele frequencies were calculated for cases and controls and compared using the Pearson's chi-squared (χ^2) test to test for association with OSCC. A *P*-value of <0.0042 (0.05/12) was used as a significance threshold for the association test to allow for multiple testing of the 12 variants present in these populations based on the Bonferroni principle. No additional correction was applied for the two populations tested. Genotypic and allelic odds ratios (ORs) with 95% confidence intervals (CIs) were calculated using the common homozygous genotype or common allele as the reference. Haplotype analysis and determination of linkage disequilibrium (LD) between variants in the same gene were performed using UNPHASED (26). SNPs that were suggestive of an allelic association (uncorrected $P < 0.05$) were further investigated for the effect of alcohol and tobacco by stratifying cases based on smoking and drinking status.

Table I. Characteristics of OSCC cases in the South African Black and Mixed Ancestry patients

	Black population	Mixed Ancestry population
Controls	*n* = 477	*n* = 427
Cases	*n* = 358	*n* = 201
Summary statistics—cases		
Age, mean years (SD)	59.8 (11.3)	60.5 (10.6)
Sex, *n* (%)		
Male	182 (50.8)	131 (65.2)
Female	176 (49.2)	70 (34.8)
Smoking status, *n* (%)		
Current smoker	100 (27.9)	131 (65.2)
Former smoker	128 (35.8)	58 (28.9)
Never smoker	130 (36.3)	10 (5.0)
Unknown	0	2 (1)
Alcohol consumption, *n* (%)		
Drinker	228 (63.7)	163 (81.1)
Non-drinker	128 (35.8)	37 (18.4)
Unknown	2 (0.6)	1 (0.5)

Results

Case–control analysis

The results for the case–control analysis in the two South African populations are shown in Table III and full genotype counts in Supplementary Table II, available at *Carcinogenesis* Online. None of the 13 variants tested were associated with OSCC in the Black South African population. The *ADH1B* 48His and *ALDH2* 504Lys alleles were absent in this population, and the *ADH7* 92Ala allele was extremely rare (only one allele observed).

In the Mixed Ancestry population, the SNP *ALDH2* +82G > A (rs886205) showed association with OSCC ($P = 0.0038$), which remained significant after accounting for multiple testing. The +82A allele had a frequency of 40.2% in cases and 48.9% in controls giving an allelic OR of 0.70 (95% CI = 0.55–0.89) and was thus associated with a reduced risk of OSCC. Suggestive associations were observed in the Mixed Ancestry population for several other polymorphisms ($P < 0.05$): *ADH1B* Arg48His ($P = 0.009$), *COX-2* −1195 A > G ($P = 0.014$), *CASP8* Asp302His ($P = 0.040$) and *MGMT* Leu84Phe ($P = 0.023$). However, these associations did not meet the required threshold for multiple testing. Variants not showing significant associations were *COX-2* −765 G > C, *ALDH2* −261 T > C, *FAS* −670 G > A,

Table II. Summary of association results in published studies

Gene	Common variant name[a]	dbSNP ID	Genetic model	OR (95% CI)	*P*-value	Cancer site	Population	Ref
ALDH2	Glu504Lys (G > A)	rs671	A versus G	1.67 (1.58–1.76)	3.27×10^{-24}	OSCC	Japanese	(7)
ALDH2	+82 A > G	rs886205	GG versus AA	4.14 (2.03–8.46)	<0.0001	OSCC	European	(14)
ALDH2	−261 C > T	rs441	CC versus TT	3.85 (1.78–8.36)	<0.0001	OSCC	European	(14)
ADH1B	Arg48His (G > A)	rs1229984	A versus G	1.79 (1.69–1.88)	7.75×10^{-24}	OSCC	Japanese	(7)
			GG + GA versus AA	0.34 (0.20–0.56)	—	OSCC	European and Latin American	(15)
ADH7	Gly92Ala	rs1573496	C versus G	0.45 (0.32–0.64)	—	OSCC	European and Latin American	(15)
			GG versus CC	0.32 (0.13–0.82)	—	Head and neck SCC	American Caucasian	(16)
FAS	−670 A > G	rs1800682	GG versus AA	1.57 (1.12–2.20)	<0.001	OSCC	Chinese	(17)
FAS	−1377 G > A	rs2234767	AA versus GG	1.62 (1.14–2.30)	<0.001	OSCC	Chinese	(17)
FASL	−844 T > C	rs763110	CC versus TT	1.72 (1.12–2.64)	<0.001	OSCC	Chinese	(17)
CASP8	Asp302His (G > C)	rs1045485	CC versus GG	0.81 (0.71–0.93)	—	Cancer meta-analysis	—	(18)
CASP8	−652 6N ins/del	rs3834129	del/del versus ins/ins	0.57 (0.36–0.88)	0.0082	Oesophageal cancer	Chinese	(19)
COX-2	−765 G > C	rs20417	GC versus GG	2.24 (1.59–3.16)	<0.0001	OSCC	Chinese	(20)
			C versus G	1.32 (0.92–1.88)	—	OSCC	Indian	(21)
COX-2	−1195 G > A	rs689466	A versus G	1.34 (1.08–1.68)	0.008	OSCC	Chinese	(20)
			A versus G	1.23 (0.80–1.87)	—	OSCC	Indian	(21)
MGMT	Leu84Phe	rs12917	FF versus LL	3.27 (1.43–7.52)	—	OSCC	European	(22)
			LF + FF versus LL	0.71 (0.51–0.98)	—	Head and neck SCC	American	(23)
			TT versus CC	1.24 (1.02–1.51)	0.035	Cancer meta-analysis	—	(24)

[a]Variant names correspond to common names used in majority of previous publications. Updated variant names based on more recent annotations of the human genome are provided in Supplementary Table I, available at *Carcinogenesis* Online.

FAS −1377 G > A, *FASL* −844T > C and *CASP8* −652 6N ins/del. The *ALDH2* 504Lys allele was absent in Mixed Ancestry subjects, as in the Black population.

Haplotype analysis

Multiple polymorphisms were genotyped in *COX-2*, *ALDH2*, *FAS* and *CASP8* genes, enabling haplotype analysis to be performed and LD to be determined (Table IV). A low level of LD was observed between variants in the different genes in both the Black and Mixed Ancestry population. The LD coefficients (r^2) for pairs of SNPs in *COX-2*, *ALDH2*, *FAS* and *CASP8* were 0.058, 0.058, 0.023 and 0.021 in the Black population controls and 0.087, 0.228 and 0.153 and 0.001 in the Mixed Ancestry population controls, respectively. Similar values were observed in OSCC cases.

In the Black population, there were no significant haplotype effects. In the Mixed Ancestry population, statistically significant haplotype associations were observed for variants in *ALDH2* ($P = 0.0028$), *FAS* ($P = 0.0031$) and *CASP8* ($P = 0.004$). However, the haplotype result observed for *ALDH2* is entirely due to the association of the *ALDH2* +82 allele observed in the single SNP analysis, with no increase in significance achieved by inclusion of the −261 variant. Haplotypes at the *FAS* gene locus were significantly associated with a reduced risk of OSCC in the Mixed Ancestry population (overall $P = 0.003$), with both −1377A and −670A alleles contributing independently to disease risk, consistent with the low level of LD between them. For *CASP8*, the haplotype −652 6Ndel:302His was significantly associated with an increased risk of OSCC ($P = 0.001$, OR = 2.37; 95% CI = 1.39–4.04), whereas the individual variants were not.

Alcohol and smoking analysis

Polymorphisms that showed a statistically significant ($P < 0.0042$) or suggestive ($P < 0.05$) association with OSCC in the case–control analysis were investigated further for gene–environment interactions. These were *ADH1B* Arg48His, *COX-2* −1195 A > G, *CASP8* Asp302His, *ALDH2* +82 G > A and *MGMT* Leu84Phe in the Mixed Ancestry population only. In this population, current, former and never-smokers account for 65.2% ($n = 131$), 28.9% (58) and 5% (10) of OSCC patients, respectively (Table I). In view of the low numbers of never-smokers, analyses were carried out only for current and former smokers, with each group being compared with controls (Table V). All SNPs analyzed showed nominal evidence of association with OSCC in current smokers ($P < 0.05$) and no associations in the smaller group of former smokers. The most significant association in current smokers compared with controls was for *MGMT* Leu84Phe ($P = 0.003$) and was more significant than the association seen in the initial case–control test for all cases combined ($P = 0.023$), with a concomitant increase in the disease risk (OR all cases 1.41 and OR current smokers 1.65).

Alcohol drinkers represent 81.1% of the Mixed Ancestry patients (Table II). Comparing drinkers to controls, all SNPs analyzed showed at least nominal evidence of association ($P < 0.05$), with the two most significant findings at *ALDH2* +82 G > A and *COX-2* −1195 A > G ($P = 0.003$ and $P = 0.004$, respectively) (Table VI), achieving greater significance in this stratified analysis compared with the initial analysis of all cases combined. Analysis of drinkers versus non-drinkers showed no significant differences between the groups as did non-drinkers versus controls, but the number of non-drinkers was small.

Discussion

In this study, we tested 13 sequence variants for association with OSCC in eight genes involved in several candidate molecular pathways, including alcohol metabolism, apoptosis, cell proliferation and DNA repair. No associations were observed in the Black South African population, whereas several significant or suggestive associations were detected in the Mixed Ancestry population. Possible explanations for the differences between these two South African populations are discussed below.

In the Mixed Ancestry population, one SNP, *ALDH2* +82 A > G (rs886205), was significantly associated with OSCC after accounting for multiple testing. Another SNP upstream of *ALDH2*, −261C > T, was not associated with OSCC in this population and the 504Lys allele was absent. ALDH2 metabolizes acetaldehyde into acetate and substitution of glutamic acid by lysine at amino acid position 504 results in a catalytically inactive subunit (27). The 504Lys allele is thought to be almost unique to Asian populations (14), and this study confirms its absence in two South African populations. The +82A allele was associated with a reduced risk of OSCC (OR = 0.70) in the Mixed Ancestry population, thus replicating the association of this SNP with OSCC previously observed in Central European populations. It should be noted that in Europeans, the +82G is the minor allele, which has an increased frequency in OSCC patients and is therefore reported as increasing cancer risk (28). Despite the difference in allele frequency between these two populations, these two observations are consistent in that the same allele (G) is increased in frequency in both OSCC populations. This SNP, located 360 bp upstream of the ATG initiation codon for *ALDH2*, is within or adjacent to known or predicted binding sites for multiple transcription factors (17,29,30). Analysis of the effect of this variant on transcriptional activity has produced conflicting results; the +82G allele has been shown to be more active than the A allele in hepatoma cells (29), but the opposite was observed in human peripheral blood leukocytes when analyzing the basal level of transcription (17). However, the latter study showed that higher levels of expression from the G allele

Table III. Association of polymorphisms with OSCC in the South African Black and Mixed Ancestry populations

Gene	SNP	Alleles (major + reference/minor)	Black population				Mixed Ancestry population			
			MAF		OR (95% CI)	*P*-value	MAF		OR (95% CI)	*P*-value*
			Cases	Controls			Cases	Controls		
ADH1B	Arg48His	G/A	0	0.000	—	—	0.054	0.098	0.52 (0.32–0.86)	0.009
ADH7	Gly92Ala	C/G	0	0.001	—	—	0.014	0.02	0.67 (0.22–2.01)	0.471
ALDH2	+82 G > A	G/A	0.247	0.252	0.98 (0.78–1.23)	0.835	0.402	0.489	0.70 (0.55–0.89)	0.004
ALDH2	−261 T > C	T/C	0.154	0.145	1.07 (0.81–1.42)	0.611	0.18	0.194	0.92 (0.67–1.25)	0.587
COX-2	−765 G > C	G/C	0.471	0.513	0.85 (0.69–1.03)	0.096	0.376	0.321	1.28 (0.99–1.64)	0.059
COX-2	−1195 A > G	A/G	0.064	0.053	1.22 (0.80–1.86)	0.343	0.103	0.155	0.63 (0.43–0.91)	0.014
MGMT	Leu84Phe	C/T	0.189	0.195	0.96 (0.75–1.24)	0.770	0.222	0.168	1.41 (1.05–1.91)	0.023
CASP8	Asp302His	G/C	0.154	0.152	1.02 (0.77–1.34)	1.000	0.169	0.126	1.42 (1.01–1.98)	0.040
CASP8	−652 6N ins/del	Ins/Del	0.518	0.502	1.06 (0.87–1.30)	0.530	0.385	0.386	0.99 (0.77–1.27)	1.000
FAS	−670 G > A	G/A	0.219	0.225	0.96 (0.76–1.22)	0.750	0.356	0.406	0.81 (0.63–1.04)	0.097
FAS	−1377 G > A	G/A	0.096	0.072	1.36 (0.95–1.94)	0.092	0.139	0.183	0.72 (0.52–1.01)	0.058
FASL	−844 T > C	T/C	0.192	0.189	1.02 (0.79–1.31)	1.000	0.416	0.386	1.13 (0.89–1.45)	0.323

*P-value corrected to three decimal places.

Table IV. Haplotype analysis for polymorphisms in *COX-2*, *FAS*, *ALDH2* and *CASP8* in the South African Black and Mixed Ancestry populations

	Black population					Mixed Ancestry population				
	Cases (%)	Controls (%)	OR (95% CI)	*P*-value	*P*-value for overall haplotype association	Cases (%)	Controls (%)	OR (95% CI)	*P*-value	*P*-value for overall haplotype association
COX-2					0.254					0.019
−765G −1195A	318 (46.4)	399 (43.7)	1.15 (0.94–1.41)	0.184		196 (51.9)	436 (52.3)	Ref	Ref	
−765C −1195A	324 (47.2)	467 (51.1)	Ref	Ref		143 (37.8)	268 (32.1)	1.19 (0.92–1.55)	0.204	
−765G −1195G	44 (6.4)	48 (5.3)	1.32 (0.86–2.04)	0.208		39 (10.3)	130 (15.6)	0.67 (0.45–0.99)	0.041	
FAS					0.238					0.003
−670G −1377G	464 (68.6)	637 (70.0)	Ref	Ref		195.6 (51.2)	339 (41.0)	Ref	Ref	
−670A −1377G	147 (21.8)	207 (22.8)	0.97 (0.76–1.24)	0.838		133.4 (34.9)	336 (40.7)	0.69 (0.53–0.90)	0.006	
−670G −1377A	65 (9.6)	66 (7.3)	1.35 (0.94–1.94)	0.104		51.38 (13.5)	151 (18.3)	0.59 (0.41–0.85)	0.004	
ALDH2					0.936					0.003
+82A −261T	166 (24.7)	231 (25.2)	0.98 (0.77–1.24)	0.872		155.3 (40.0)	405 (48.7)	Ref	Ref	
+82G −261T	403 (6.0)	550 (60.0)	Ref	Ref		162.7 (41.9)	266 (32.0)	1.60 (1.22–2.09)	0.001	
+82G −261C	103 (15.3)	135 (14.7)	1.04 (0.78–1.39)	0.782		68.32 (17.6)	161 (19.4)	1.11 (0.79–1.55)	0.559	
CASP8					0.406					0.004
−652 6N Ins 302D	256 (39.7)	361.2 (39.9)	0.99 (0.79–1.26)	0.959		200.2 (54.4)	436 (53.0)	Ref	Ref	
−652 6N Del 302D	290 (45.0)	406.8 (44.9)	Ref	Ref		102.8 (27.9)	281 (34.2)	0.80 (0.59–1.07)	0.126	
−652 6N Ins 302H	52.04 (8.1)	92.78 (10.2)	0.79 (0.53–1.18)	0.238		25.8 (7.0)	69 (8.4)	0.81 (0.48–1.39)	0.448	
−652 6N Del 302H	45.96 (7.1)	45.22 (5.0)	1.43 (0.83–2.44)	0.199		39.2 (10.7)	36 (4.4)	2.37 (1.39–4.04)	0.001	

could be induced by acetaldehyde or ethanol, suggesting that this SNP may contribute to interindividual or allelic differences in acetaldehyde elimination (17). The association of *ALDH2* +82A > G in the South African Mixed Ancestry population is consistent with many other studies which have found associations of other variants at this gene with oesophageal cancer, including patients of European, Chinese and Japanese descent (8,9, 28,31–33). Interestingly, this is a region of extended LD in European populations, which contains other plausible candidate genes (9).

The SNPs *ADH1B* Arg48His, *COX-2* −1195A > G, *CASP8* Asp302His and *MGMT* Leu84Phe showed some evidence of association with OSCC in the Mixed Ancestry population with $P < 0.05$ but did not survive the Bonferroni threshold which was set for multiple testing. However, given that this threshold is somewhat conservative and that there is prior evidence for the association of these variants with OSCC or head and neck squamous cell carcinoma in other populations, we have referred to these as suggestive associations with OSCC which require future follow-up in an expanded sample.

The associations of haplotypes at the *FAS* gene locus which contained either the −1377A or −670A alleles with a reduced risk of OSCC in the Mixed Ancestry population are only partially consistent with data from the Han Chinese population, in which −670A was associated with a reduced risk of OSCC but −1377A had the opposite effect (34). The *FAS* protein and its ligand, *FASL*, play a key role in the induction of apoptosis, and *FAS* expression is found to be reduced in OSCC tissues (35,36). However, a recent study by Chen *et al.* (37) showed that constitutive expression of *FAS* is required for optimal growth of tumors and complete loss of *FAS* is rarely observed. Functional studies have shown that *FAS* −1377A and −670G disrupt the Sp1 and STAT1 transcription factor-binding sites, respectively, which both lead to reduced expression of *FAS* (15,38,39). Taken together, these data suggest that the *FAS* −1377A and −670G alleles would both be associated with an increased risk of cancer, provided that *FAS* is constitutively expressed, but population-specific effects may occur.

In our study, neither of the *CASP8* variants, Asp302His and -652 6N ins/del, were significantly associated with OSCC. However, in the haplotype analysis, the *CASP8* −652 6Ndel:302His haplotype was significantly associated with an increased risk OSCC. Previous independent association studies of *CASP8* variants with cancer have produced somewhat conflicting results. The −652 6N deletion allele has been associated with reduced susceptibility to multiple cancers, including oesophageal cancer, in the Chinese population (25), but this was not replicated for breast, colorectal and prostate cancer in several other populations (40). A meta-analysis of 55 studies reported a reduced overall risk of cancer for the −652 6N deletion and 302His alleles, although stratified analysis suggested a reduced risk for estrogen-related cancers but an increased risk for brain tumors (18). It is possible that population-specific or cancer subtype-specific effects may occur, and results may also be influenced by differing environmental triggers.

Gene–environment interactions are known to exist in susceptibility to OSCC, particularly those involving an interaction between alcohol metabolism genes and alcohol intake or smoking status. The risk of disease conferred by the *ALDH2* 504Lys allele, for example, is related to the amount of alcohol consumed (8, 41), but association studies in non-drinkers have reported conflicting results (8, 32, 41). Smoking status also appears to interact with this variant, with smokers having a greater risk of disease than non-smokers (8). The association observed in this study for *ALDH2* +82A > G in the South African Mixed Ancestry population was somewhat stronger when the data were stratified to include only those cases that drink. This follows a similar trend to that observed by Hashibe *et al.* (28), who found an increased disease risk for A/G heterozygotes and G/G homozygotes in both light and medium/heavy drinkers in a European population. Gene–environmental interactions involving *ADH1B* His48Arg have also been conflicting. Cui *et al.* (8) found no interaction of this SNP with drinking and smoking status in the Japanese population. This is in contrast to a study of upper aerodigestive cancers,

Table V. Analysis of the effect of smoking for polymorphisms that show a significant or suggestive association with OSCC in the Mixed Ancestry population

Gene	Current smokers (n = 131) versus controls (n = 427)			Former smokers (n = 58) versus controls (n = 427)		
	MAF: controls/ current smokers	OR (95% CI)	P-value	MAF: controls/ former smokers	OR (95% CI)	P-value
ADH1B Arg48His	0.098/0.047	0.45 (0.24–0.84)	0.011	0.098/0.096	0.98 (0.51–1.91)	1.000
COX-2 −1195 A > G	0.155/0.100	0.61 (0.38–0.95)	0.029	0.155/0.129	0.81 (0.46–1.44)	0.468
CASP8 Asp302His	0.126/0.180	1.53 (1.04–2.24)	0.028	0.126/0.179	1.51 (0.90–2.56)	0.119
ALDH2 +82 G > A	0.489/0.418	0.75 (0.56–0.99)	0.045	0.489/0.407	0.72 (0.48–1.06)	0.092
MGMT Leu84Phe	0.168/0.250	1.65 (1.18–2.32)	0.003	0.168/0.186	1.14 (0.69–1.87)	0.615

Table VI. Analysis of the effect of alcohol consumption for polymorphisms that show a significant or suggestive association with OSCC in the Mixed Ancestry population

Gene	Case-only analysis: drinkers (n = 163) versus non-drinkers (n = 37)			Drinkers (n = 163) versus control (n = 427)			Non-drinkers (n = 37) versus controls (n = 427)		
	MAF: non-drinkers/ drinkers	OR (95% CI)	P-value	MAF: controls/ drinkers	OR (95% CI)	P-value	MAF: controls/ non-drinkers	OR (95% CI)	P-value
ADH1B Arg48His	0.042/0.057	1.38 (0.40–4.82)	0.612	0.098/0.057	0.55 (0.33–0.94)	0.026	0.098/0.042	0.40 (0.12–1.30)	0.116
COX-2 −1195 A > G	0.162/0.090	0.51 (0.25–1.06)	0.066	0.155/0.090	0.54 (0.35–0.83)	0.004	0.155/0.162	1.05 (0.55–2.01)	0.873
CASP8 Asp302His	0.111/0.184	1.80 (0.82–3.97)	0.139	0.126/0.184	1.57 (1.10–2.23)	0.012	0.126/0.111	0.87 (0.41–1.87)	0.721
ALDH2 +82 G > A	0.432/0.394	0.85 (0.51–1.42)	0.541	0.489/0.394	0.68 (0.52–0.88)	0.003	0.489/0.432	0.79 (0.49–1.28)	0.347
MGMT Leu84Phe	0.176/0.234	1.43 (0.75–2.76)	0.277	0.168/0.234	1.52 (1.10–2.08)	0.010	0.168/0.176	1.06 (0.57–1.97)	0.863

including OSCC, in Central European countries where only drinkers and smokers showed an altered susceptibility to disease for this genotype (42). In a study in the Chinese population, *ADH1B* His48Arg was not associated with oesophageal cancer in drinkers or non-drinkers (32). Despite this lack of consistency, one study has shown that the combination of risk SNPs for *ALDH2* Glu504Lys and *ADH1B* His48Arg, together with drinking alcohol and smoking, increases the risk of OSCC synergistically compared with individuals with no risk genotypes who refrain from alcohol consumption and smoking [OR = 189.26 (95% CI = 95.11–376.63)] (8).

In our study, all the significant or suggestive associations detected in cases from the Mixed Ancestry population were also observed in alcohol drinkers but not in non-drinkers. However, the proportion of non-drinkers in this population was very small, with low power to detect any association. Interestingly, the protective effect of the *COX-2* −1195 A > G SNP (P = 0.004) was strengthened in the drinkers subgroup relative to the whole case sample (P = 0.014), again suggestive of a possible gene–environmental interaction, but this requires confirmation by analysis of a larger sample of non-drinking OSCC cases. In relation to smoking status, again all associations observed in the full sample of OSCC cases were also detected in smokers but not in former smokers or non-smokers (those who have never smoked). However, former smokers and non-smokers were a minority in this population, so there was low power to detect associations in these subgroups. The stronger association observed for *MGMT* Leu84Phe with OSCC in smoking cases versus controls (P = 0.003) compared with all cases versus controls (P = 0.023) suggests a possible interaction between smoking and the *MGMT* DNA repair pathway but requires confirmation by analysis of a larger sample of non-smoking cases.

The differences in the results of the genetic association tests between the Black and Mixed Ancestry populations in this study are striking. A number of associations reported previously in European and Asian populations were also detected in the Mixed Ancestry population, albeit with varying levels of support. However, none of the variants tested showed any evidence of association in the Black population. There are several possible explanations for these differences. One is the absence or extreme rarity of the risk SNP in this population, such as was the case for *ADH1B* 48His, *ALDH2* 504Lys and *ADH7* 92Ala. Another is a lack of statistical power to detect modest genetic effects in the samples available. The sample size was larger in the Black population than in the Mixed Ancestry (N = 835 and 628, respectively), and, in several instances, there was high power to detect the same effect seen in the Mixed Ancestry population. However, power in the Black population sample varied substantially, depending on the size of the effect (OR) and the minor allele frequency. Thus, for an SNP with a MAF of 25% and an allelic OR of 1.5, we had 96% power to detect association (at significance level 0.05). However, for *COX-2* −1195A > G (OR 1.34), for example, the MAF was 5% and power was only 30%, whereas for *ALDH2* +82G > A (MAF = 25%, OR = 1.43), the power to detect the same effect seen in the Mixed Ancestry population was >90%. Another possibility is that the SNPs associated with OSCC in other populations may not be the true causal variants and may be merely tagging SNPs, which are in high LD with the causal variant. There is generally a lower level of LD in African populations, and hence, associations may not be detected if the causal variant is not genotyped (43).

A further and potentially important explanation for these results is differences in environmental exposures between these two populations. The significant or suggestive association with OSCC of variants in genes involved in alcohol and acetaldehyde metabolism (*ALDH2* and *ADH1B*) and DNA repair of alkylation-induced mutagenesis (*MGMT*) in the Mixed Ancestry population is consistent with the high rates of smoking and alcohol use in this cohort. However, the proportion of smokers and drinkers was lower in the Black OSCC patients, thus a genetic interaction between, for example, a genetic variant in an aldehyde dehydrogenase gene and alcohol consumption or smoking might be weaker or absent in this population and the genetic association would not be observed. Epidemiological studies in Black South Africans have investigated the contribution of alcohol consumption and smoking to oesophageal cancer risk (44–46), the most recent of which suggests that while alcohol use in itself does

not increase risk (OR 0.9), smoking does (OR 1.9) and the combination of smoking and drinking increases risk further (OR 4.4) (46). However, a major risk factor in this population is residence in the Transkei region of the Eastern Cape Province (47,48), with long-term residence of 35 years or more associated with high risk (OR 14.7) in females (46). Nutritional factors such as deficiency in micronutrients (49,50), and the consumption of maze contaminated with *Fusarium* species, which produce mycotoxins (51) have been proposed as risk factors. Taken together, these studies suggest that dietary factors are likely to be an important component of risk for OSCC in this population. If so, the nature of interacting genetic factors may be substantially different to those found in the Mixed Ancestry and other populations. In view of the success of genome-wide association scans in defining the genetic components of complex disease, this approach is likely to be the most effective means of identifying the relevant genetic pathways involved in the pathogenesis of OSCC in these two populations once sufficient sample sizes for genome-wide studies are available. The identification of such pathways may shed further light on the role of dietary factors in OSCC.

In conclusion, this study provides evidence that several genetic variants in genes involved in alcohol metabolism and DNA repair contribute to genetic susceptibility to OSCC in the Mixed Ancestry but not in the Black population of South Africa. This may be explained both by differences in the genetic history and architecture of these populations and by different environmental exposures.

Supplementary material

Supplementary Tables I and II can be found at http://carcin.oxfordjournals.org/

Funding

Association for International Cancer Research (grant number 09-0625), the Medical Research Council UK, The Generation Trust, the National Institutes of Health Research Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation, the International Centre for Genetic Engineering and Biotechnology (ICGEB), the South African Medical Research Council and the University of Cape Town to M.I.P. National Institutes of Health, USA (06565) and the South African Medical Research Council to E.R.

Acknowledgements

We wish to thank Antoinette Olivier and Zenaria Abbas for assisting with the sample collection and processing and the patients and healthy controls for their participation in this study.

Conflict of Interest Statement: None declared.

References

1. Parkin,D.M. *et al*. (2005) Global cancer statistics, 2002. *CA Cancer J. Clin.*, **55**, 74–108.
2. Hendricks,D. *et al*. (2002) Oesophageal cancer in Africa. *IUBMB Life*, **53**, 263–268.
3. Lam,A.K. (2000) Molecular biology of esophageal squamous cell carcinoma. *Crit. Rev. Oncol. Hematol.*, **33**, 71–90.
4. Somdyala,N.I. *et al*. (2010) Cancer incidence in a rural population of South Africa, 1998–2002. *Int. J. Cancer*, **127**, 2420–2429.
5. Lao-Sirieix,P. *et al*. (2010) Genetic predisposition to gastro-oesophageal cancer. *Curr. Opin. Genet. Dev.*, **20**, 210–217.
6. Wang,L.D. *et al*. (2010) Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nat. Genet.*, **42**, 759–763.
7. Abnet,C.C. *et al*. (2010) A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat. Genet.*, **42**, 764–767.
8. Cui,R. *et al*. (2009) Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology*, **137**, 1768–1775.
9. McKay,J.D. *et al*. (2011) A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet.*, **7**, e1001333.
10. Li,D. *et al*. (2010) The 341C/T polymorphism in the GSTP1 gene is associated with increased risk of oesophageal cancer. *BMC Genet.*, **11**, 47.
11. Li,D. *et al*. (2005) Association of cytochrome P450 2E1 genetic polymorphisms with squamous cell carcinoma of the oesophagus. *Clin. Chem. Lab. Med.*, **43**, 370–375.
12. Dandara,C. *et al*. (2006) Gene-environment interaction: the role of SULT1A1 and CYP3A5 polymorphisms as risk modifiers for squamous cell carcinoma of the oesophagus. *Carcinogenesis*, **27**, 791–797.
13. de Wit,E. *et al*. (2010) Genome-wide analysis of the structure of the South African coloured population in the Western Cape. *Hum. Genet.*, **128**, 145–153.
14. Li,H. *et al*. (2009) Refined geographic distribution of the oriental ALDH2*504Lys (nee 487Lys) variant. *Ann. Hum. Genet.*, **73**, 335–345.
15. Huang,Q.R. *et al*. (1997) Identification and characterization of polymorphisms in the promoter region of the human Apo-1/Fas (CD95) gene. *Mol. Immunol.*, **34**, 577–582.
16. Wei,S. *et al*. (2010) A single nucleotide polymorphism in the alcohol dehydrogenase 7 gene (alanine to glycine substitution at amino acid 92) is associated with the risk of squamous cell carcinoma of the head and neck. *Cancer*, **116**, 2984–2992.
17. Kimura,Y. *et al*. (2009) A promoter polymorphism in the ALDH2 gene affects its basal and acetaldehyde/ethanol-induced gene expression in human peripheral blood leukocytes and HepG2 cells. *Alcohol Alcohol.*, **44**, 261–266.
18. Yin,M. *et al*. (2010) CASP8 polymorphisms contribute to cancer susceptibility: evidence from a meta-analysis of 23 publications with 55 individual studies. *Carcinogenesis*, **31**, 850–857.
19. Morin,P.A. *et al*. (1999) High-throughput single nucleotide polymorphism genotyping by fluorescent 5' exonuclease assay. *Biotechniques*, **27**, 538–552.
20. Zhang,X. *et al*. (2005) Identification of functional genetic variants in cyclooxygenase-2 and their association with risk of esophageal cancer. *Gastroenterology*, **129**, 565–576.
21. Upadhyay,R. *et al*. (2009) Functional polymorphisms of cyclooxygenase-2 (COX-2) gene and risk for esophageal squmaous cell carcinoma. *Mutat. Res.*, **663**, 52–59.
22. Hall,J. *et al*. (2007) The association of sequence variants in DNA repair and cell cycle genes with cancers of the upper aerodigestive tract. *Carcinogenesis*, **28**, 665–671.
23. Huang,W.Y. *et al*. (2005) Selected genetic polymorphisms in MGMT, XRCC1, XPD, and XRCC3 and risk of head and neck cancer: a pooled analysis. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 1747–1753.
24. Zhong,Y. *et al*. (2010) Effects of O6-methylguanine-DNA methyltransferase (MGMT) polymorphisms on cancer: a meta-analysis. *Mutagenesis*, **25**, 83–95.
25. Sun,T. *et al*. (2007) A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. *Nat. Genet.*, **39**, 605–613.
26. Dudbridge,F. (2008) Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum. Hered.*, **66**, 87–98.
27. Crabb,D.W. *et al*. (1989) Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. *J. Clin. Invest.*, **83**, 314–316.
28. Hashibe,M. *et al*. (2006) Evidence for an important role of alcohol- and aldehyde-metabolizing genes in cancers of the upper aerodigestive tract. *Cancer Epidemiol. Biomarkers Prev.*, **15**, 696–703.
29. Chou,W.Y. *et al*. (1999) An A/G polymorphism in the promoter of mitochondrial aldehyde dehydrogenase (ALDH2): effects of the sequence variant on transcription factor binding and promoter strength. *Alcohol Clin. Exp. Res.*, **23**, 963–968.
30. Stewart,M.J. *et al*. (1998) Binding and activation of the human aldehyde dehydrogenase 2 promoter by hepatocyte nuclear factor 4. *Biochim. Biophys. Acta*, **1399**, 181–186.
31. Ding,J.H. *et al*. (2009) Polymorphisms of alcohol dehydrogenase-2 and aldehyde dehydrogenase-2 and esophageal cancer risk in Southeast Chinese males. *World J. Gastroenterol.*, **15**, 2395–2400.
32. Ding,J.H. *et al*. (2009) Alcohol dehydrogenase-2 and aldehyde dehydrogenase-2 genotypes, alcohol drinking and the risk for esophageal cancer in a Chinese population. *J. Hum. Genet.*, **55**, 97–102.
33. Yokoyama,A. *et al*. (2002) Genetic polymorphisms of alcohol and aldehyde dehydrogenases and glutathione S-transferase M1 and drinking, smoking, and diet in Japanese men with esophageal squamous cell carcinoma. *Carcinogenesis*, **23**, 1851–1859.

34. Sun,T. *et al*. (2004) Polymorphisms of death pathway genes FAS and FASL in esophageal squamous-cell carcinoma. *J. Natl Cancer Inst.*, **96**, 1030–1036.
35. Kase,S. *et al*. (2002) Expression of Fas and Fas ligand in esophageal tissue mucosa and carcinomas. *Int. J. Oncol.*, **20**, 291–297.
36. Gratas,C. *et al*. (1998) Up-regulation of Fas (APO-1/CD95) ligand and down-regulation of Fas expression in human esophageal cancer. *Cancer Res.*, **58**, 2057–2062.
37. Chen,L. *et al*. (2010) CD95 promotes tumour growth. *Nature*, **465**, 492–496.
38. Kanemitsu,S. *et al*. (2002) A functional polymorphism in fas (CD95/APO-1) gene promoter associated with systemic lupus erythematosus. *J. Rheumatol.*, **29**, 1183–1188.
39. Sibley,K. *et al*. (2003) Functional FAS promoter polymorphisms are associated with increased risk of acute myeloid leukemia. *Cancer Res.*, **63**, 4327–4330.
40. Haiman,C.A. *et al*. (2008) A promoter polymorphism in the CASP8 gene is not associated with cancer risk. *Nat. Genet.*, **40**, 259–260; author Reply, 260–261.
41. Lewis,S.J. *et al*. (2005) Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 1967–1971.
42. Hashibe,M. *et al*. (2008) Multiple ADH genes are associated with upper aerodigestive cancers. *Nat. Genet.*, **40**, 707–709.
43. Teo,Y.Y. *et al*. (2010) Methodological challenges of genome-wide association analysis in Africa. *Nat. Rev. Genet.*, **11**, 149–160.
44. McGlashan,N.D. *et al*. (1982) Cancer of the oesophagus and the use of tobacco and alcoholic beverages in Transkei, 1975-6. *Int J Cancer*, **29**, 249–256.
45. Segal,I. *et al*. (1988) Factors associated with oesophageal cancer in Soweto, South Africa. *Br. J. Cancer*, **58**, 681–686.
46. Pacella-Norman,R. *et al*. (2002) Risk factors for oesophageal, lung, oral and laryngeal cancers in black South Africans. *Br. J. Cancer*, **86**, 1751–1756.
47. Rose,E.F. (1973) Esophageal cancer in the Transkei: 1955-69. *J. Natl Cancer Inst.*, **51**, 7–16.
48. Burrell,R.J. (1962) Esophageal cancer among Bantu in the Transkei. *J. Natl Cancer Inst.*, **28**, 495–514.
49. Van Rensburg,S.J. (1981) Epidemiologic and dietary evidence for a specific nutritional predisposition to esophageal cancer. *J. Natl Cancer Inst.*, **67**, 243–251.
50. Van Rensburg,S.J. *et al*. (1983) Nutritional status of African populations predisposed to esophageal cancer. *Nutr. Cancer*, **4**, 206–216.
51. Shephard,G.S. *et al*. (2007) Exposure assessment for fumonisins in the former Transkei region of South Africa. *Food Addit. Contam.*, **24**, 621–629.

Received July 27, 2011; revised August 25, 2011; accepted September 10, 2011