Research Paper

Association of *ALDH3B2* gene polymorphism and risk factors with susceptibility of esophageal squamous cell carcinoma in a Chinese population: a case-control study involving 2,358 subjects

Jun Yin[1,2,*], Weifeng Tang[1,*], Tao Long[1,*], Huiwen Pan[1], Jianchao Liu[1], Lu Lv[1], Chao Liu[1], Yijun Shi[1], Jingfeng Zhu[1], Yangyong Sun[1], Aizhong Shao[1], Qiang Zhou[1], Zhengbing Ren[1], Guowen Ding[1], Suocheng Chen[1], Yan Liu[3], Jun Yao[4], Hao Ding[5], Yulan Yan[5], Haiyong Gu[6], Cheng Qian[2], Liming Wang[7], Qun Wang[2] and Lijie Tan[2]

[1]Department of Cardiothoracic Surgery, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu, 212002, China

[2]Department of Thoracic Surgery, Zhongshan Hospital of Fudan University, Shanghai, 200032, China

[3]Genesky Biotechnologies Inc., Shanghai, 201315, China

[4]Department of Gastroenterology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu, 212002, China

[5]Department of Respirology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu, 212002, China

[6]Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiaotong University, Shanghai, 200030, China

[7]Cancer Institute, Department of Chemotherapy, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu, 212002, China

[*]These authors contributed equally to this work

Correspondence to: Liming Wang, **email:** yin912@outlook.com
          Qun Wang, **email:** wang.qun@zs-hospital.sh.cn
          Lijie Tan, **email:** tan.lijie@zs-hospital.sh.cn

ABSTRACT

Background: Esophageal cancer (EC) is the sixth leading cause of cancer-associated death worldwide. The interaction of environmental risk factors and genetic factors might contribute to the carcinogenesis of EC synergistically.

Results: All seven single locus polymorphisms of ALDH3B2 were not associated with risk of ESCC as evaluated by allelic, dominant, co-dominant, recessive and Cochran-Armitage trend tests. Stratified analyses showed these SNPs were not correlated with the susceptibility of ESCC according to different age, gender, cigarette smoking and alcohol drinking status. None of the major haplotypes were related with ESCC susceptibility.

Materials and Methods: We conducted a hospital-based case–control study to evaluate the combined effects of environmental risk factors and the single nucleotide polymorphisms (SNPs) of ALDH3B2 gene on the development of esophageal squamous carcinoma (ESCC). A total of 1043 ESCC cases and 1315 controls were recruited for this study. Seven ALDH3B2 SNPs and four environmental factors were selected as independent variables. ALDH3B2 SNPs were determined by ligation detection reaction method.

Conclusions: Our study suggested that ALDH3B2 rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 polymorphisms were not implicated with altered susceptibility of ESCC according to different age, gender, cigarette smoking and alcohol drinking status. Yet this conclusion needs to be verified in larger studies among different ethnic populations with validation design, the biological function of these SNPs in carcinogenesis are subject to further investigation.

INTRODUCTION

Esophageal cancer ranks the ninth most common cancer and the sixth most common cause of cancer death worldwide [1]. Despite remarkable advances in the therapeutic strategy, extensive treatment may be associated with a noticeable decline in health-related quality of life and yet a poor prognosis [2]. Approximately 70% of global esophageal cancer cases occur in China, with esophageal squamous cell carcinoma (ESCC) being the histopathological form in the vast majority of cases (> 90%) [3]. Alcohol consumption [4, 5], tobacco use [4, 6], poor oral hygiene, low socioeconomic status and nutritional deficiencies have been reported risk factors for esophageal cancer [7–10]. The fact that only a subset of cohort that are exposed to the risk factors eventually develop esophageal cancer suggested a critical role of genetic factors, including single nucleotide polymorphisms (SNPs), in the esophageal carcinogenesis.

Among all risk factors, others and we have repeatedly verified the indisputable role of alcohol consumption in the ESCC carcinogenesis [11, 12]. Extensive evaluations and reviews of alcohol-related cancers have been published in the Monographs of the International Agency of Research on Cancer (IARC) [13, 14], with most convincing evidence (Group I) for alcohol-drinking-related cancers is targeted on esophagus and some other organs [15]. Specific enzymes, whose activity and expression are influenced by genetic polymorphisms, regulate the metabolism of alcohol [11]. In humans, alcohol is primarily metabolized by two major groups of enzymes termed alcohol dehydrogenases (ADHs) and aldehyde dehydrogenases (ALDHs) [16]. In the cytosol of hepatocytes, ADHs catalyse the oxidation of ethanol to acetaldehyde, which is further oxidized to acetate by ALDHs in the mitochondria [11]. Although alcohol is not a carcinogen *per se*, its metabolite acetaldehyde is a toxin and carcinogen that rapidly binds to protein and DNA. It has profound effects on carcinogenesis by forming with DNA carcinogenic DNA adducts, by inhibiting DNA repair and by regulating DNA methylation. Acetaldehyde is degradated by ALDHs, which renders ALDHs a pivotal role in the carcinogenesis. Indeed, high ALDH1 expression predicts unfavorable outcomes in patients with ESCC [17]. Individuals with *ALDH2* Lys allele possess a higher risk of esophageal cancer, in correlation with a higher concentration of blood acetaldehyde after drinking alcohol [18]. There is a strong association between *ALDH2* Glu487Lys polymorphism and the risk of esophageal cancer [19]. *ALDH2* rs671 [20] and rs886205 [21] polymorphisms have also been demonstrated to correlate with ESCC, respectively.

Similarly, the aldehyde dehydrogenase 3 family member B2 (ALDH3B2) is also a key member of ALDH family. Originally identified as ALDH8, ALDH3B2 encodes a member of the aldehyde dehydrogenase family, a group of isozymes that may play a major role in the detoxification of aldehydes generated by alcohol metabolism and lipid peroxidation. As compared with ALDH1/2, little is known on ALDH3 family (including ALDH3B2) with respect to their roles in carcinogenesis. The association between ALDH3B2 polymorphisms and ESCC has not been investigated. Hence, in this hospital-based case-control study, we performed genotyping analyses of the seven SNPs in 1043 ESCC cases and 1315 controls in a Chinese population.

RESULTS

Characteristics of the study subjects

The characteristics of the study subjects, including the demographics and environmental risk factors, are presented in Table 1. The controls and cases were well matched in age and gender (χ^2 test, $p = 0.121$ and 0.880, respectively). However, the cigarette-smoking rate (43.53% vs.26.70%, $p < 0.001$) and alcohol drinking rate (31.54% vs.7.07%, $p < 0.001$) were both significantly higher in the ESCC cases.

As shown in Table 2, the genotyping successful rates were all beyond 98.81%. In the control subjects, the genotype frequencies for the seven polymorphisms reached Hardy-Weinberg equilibrium (*p-value* for HWE, all $p > 0.05$, Table 2).

Associations between risk of ESCC and seven polymorphisms

As shown in Table 3, the single locus analyses showed no statistically significant difference in genotype frequencies of seven SNPs between the cases and controls ($p > 0.05$). As assessed by the allelic, dominant, co-dominant, recessive and Cochran-Armitage trend tests, there are no correlations between these seven polymorphism sites with the risk of ESCC (Table 3).

Stratification analyses on seven polymorphisms and risk of ESCC

To further evaluate the effects of *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 on ESCC risk with different gender, age, smoking and alcohol drinking status, stratification analyses were performed as demonstrated in the Tables 4–10.

Our analyses demonstrated that neither gender, age, cigarette smoking nor alcohol drinking has detectable impacts on the susceptibility of ESCC after stratified analyses.

Table 1: Distribution of selected demographic variables and risk factors in ESCC cases and controls

Variable	Cases (*n* = 1043)		Controls (*n* = 1315)		*p*
	n	*%*	*n*	*%*	
Age (years) mean ± SD	63.07(± 7.27)		62.88 (± 9.74)		0.607
Age (years)					
< 63	471	45.16	636	48.37	
≥ 63	572	54.84	679	51.63	0.121
Sex					
Male	758	72.67	952	72.40	
Female	285	27.33	363	27.63	0.880
Tobacco use					
Never	589	56.47	964	73.30	
Ever	454	43.53	351	26.70	**< 0.001**
Alcohol use					
Never	714	68.46	1222	92.93	
Ever	329	31.54	93	7.07	**< 0.001**

[a] Two-sided χ^2 test and student t test; Bold values are statistically significant (*p* < 0.05).

Table 2: Primary information for *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907, rs9787887 polymorphisms

Genotyped SNPs	rs34589365	rs3741172	rs4646823	rs78402723	rs7947978	rs866907	rs9787887
Ancestral Allele	A	G	G	G	C	A	G
Chromosome	11	11	11	11	11	11	11
Gene (ID)	ALDH3B2(222)	ALDH3B2(222)	ALDH3B2(222)	ALDH3B2(222)	ALDH3B2(222)	ALDH3B2(222)	ALDH3B2(222)
Function	intron-variant	reference, synonymous-codon	utr-variant-5-prime	utr-variant-3-prime	intron-variant, utr-variant-5-prime	utr-variant-3-prime	intron-variant
Chr Pos (Genome Build 38.p7)	67668505	67663227	67666945	67663118	67674572	67663011	67673240
Regulome DB Score[a]	5	5	4	4	4	4	5
TFBS[b]	Y	--	Y	--	--	--	--
nsSNP	--	--	--	--	--	--	--
MAF[c] for Chinese in database	0.183	--	0.098	0.058	0.098	0.128	--
MAF in our controls (*n* = 1315)	0.126	0.084	0.105	0.084	0.105	0.105	0.335
p value for HWE[d] test in our controls	0.110	0.784	0.294	0.808	0.279	0.294	0.258
Genotyping method[e]	LDR	LDR	LDR	LDR	LDR	LDR	LDR
% Genotyping value	99.02%	99.02%	99.02%	98.98%	99.02%	99.02%	98.81%

[a]http://www.regulomedb.org/;
[b]TFBS: Transcription Factor Binding Site (https://snpinfo.niehs.nih.gov/cgi-bin/snpinfo/snpfunc.cgi);
[c]MAF: minor allele frequency;
[d]HWE: Hardy–Weinberg equilibrium;
[e]LDR: ligation detection reaction.

Table 3: Main effects of *ALDH3B2* SNPs on ESCC risk

Locus	Geno-type	Case	Control	Co-dominant χ² P	Dominant χ² P	Recessive χ² P	Cochran-Armitage trend χ² P	Allelic test χ² P
rs34589365 A/G	AA	13 (1.27%)	27 (2.06%)					
	AG	228 (22.24%)	275 (20.99%)	2.544; 0.2803	0.0678; 0.7945	2.147; 0.1429	0.0290; 0.8649	0.0294; 0.8639
	GG	784 (76.49%)	1008 (76.95%)					
rs3741172 G/A	GG	839 (81.85%)	1100 (83.97%)					
	AG	176 (17.17%)	200 (15.27%)	1.906; 0.3855	1.828; 0.1764	0.3051; 0.5807	1.906; 0.1674	1.92; 0.1658
	AA	10 (0.98%)	10 (0.76%)					
rs4646823 G/T	GG	15 (1.46%)	18 (1.37%)					
	GT	202 (19.71%)	239 (18.24%)	0.8581; 0.6511	0.8566; 0.3547	0.0330; 0.8559	0.7819; 0.3766	0.8008; 0.3709
	TT	808 (78.83%)	1053 (80.38%)					
rs78402723 G/A	GG	849 (82.91%)	1099 (83.89%)					
	AG	166 (16.21%)	201 (15.34%)	0.4358; 0.8042	0.4023; 0.5259	0.0950; 0.7579	0.4343; 0.5099	0.4377; 0.5082
	AA	9 (0.88%)	10 (0.76%)					
rs7947978 C/A	CC	804 (78.44%)	1054 (80.46%)					
	CA	209 (20.39%)	238 (18.17%)	1.963; 0.3747	1.442; 0.2298	0.1874; 0.6651	0.9688; 0.325	0.9794; 0.3223
	AA	12 (1.17%)	18 (1.37%)					
rs866907 A/G	AA	12 (1.17%)	18 (1.37%)					
	AG	204 (19.90%)	239 (18.24%)	1.171; 0.5568	0.7535; 0.3854	0.1874; 0.6651	0.4626; 0.4964	0.4686; 0.4936
	GG	809 (78.93%)	1053 (80.38%)					
rs9787887 G/A	GG	129 (12.63%)	156 (11.92%)					
	AG	454 (44.47%)	565 (43.16%)	0.9959; 0.6078	0.9504; 0.3296	0.2748; 0.6001	0.9316; 0.3344	0.9567; 0.328
	AA	438 (42.90%)	588 (44.92%)					

Table 4: Stratified analyses between *ALDH3B2* rs34589365 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control) [a] AA	GG	AG	AG+GG	AA	Adjusted OR [b] (95%CI); p GG	AG	AG+GG	GG vs. (AA+AG)
Sex									
Male	12/19	554/717	179/212	733/929	1.00	0.817 (0.39–1.70); p:0.588	0.748 (0.35–1.58); p:0.446	0.800 (0.39–1.66); p:0.547	0.934 (0.75–1.17); p:0.549
Female	1/8	230/291	49/63	279/354	1.00	1.158 (0.02–1.27); p:0.085	0.161 (0.02–1.33); p:0.079	0.159 (0.02–1.28); p:0.085	1.122 (0.75–1.68); p:0.573
Age									
< 63	5/14	347/478	114/140	461/618	1.00	0.492 (0.18–1.38); p:0.239	0.439 (0.15–1.25); p:0.151	0.479 (0.17–1.34); p:0.168	0.939 (0.71–1.24); p:0.658
≥ 63	8/13	437/530	114/135	551/665	1.00	0.746 (0.31–1.82); p:0.659	0.729 (0.29–1.82); p:0.497	0.746 (0.31–1.82); p:0.518	1.000 (0.76–1.31); p:0.999
Smoking status									
Never	8/21	449/743	117/196	566/939	1.00	0.630 (0.28–1.44); p:0.268	0.638 (0.27–1.49); p:0.295	0.632 (0.28–1.44); p:0.269	1.049 (0.82–1.35); p:0.706
Ever	5/6	335/265	111/79	446/344	1.00	0.659 (0.20–2.18); p:0.550	0.593 (0.18–2.01); p:0.533	0.643 (0.20–2.12); p:0.547	0.926 (0.67–1.28); p:0.642
Alcohol consumption									
Never	10/26	555/945	138/246	693/1191	1.00	0.655 (0.31–1.37); p:0.257	0.686 (0.32–1.46); p:0.327	0.661 (0.32–1.38); p:0.267	1.079 (0.86–1.35); p:0.508
Ever	3/1	229/63	90/29	319/92	1.00	0.825 (0.08–8.07); p:1.000	0.967 (0.10–9.66); p:1.000	0.865 (0.09–8.42); p:1.000	1.173 (0.71–1.93); p:0.530

[a]The genotyping success rate was 99.02% for rs34589365; [b]Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n \leq 5$; Bonferroni correction was performed to correct the p value, bold values are statistically significant ($p < 0.0125$).

Table 5: Stratified analyses between *ALDH3B2* rs3741172 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control)[a]				Adjusted OR b (95%CI); *p*				
	GG	AA	AG	AA+AG	GG	AA	AG	AA+AG	AA vs. (GG+AG)
Sex									
Male	622/798	5/7	118/143	123/150	1.00	1.091 (0.35–3.46); *p*:1.000	0.945 (0.72–1.23); *p*:0.674	0.951 (0.73–1.23); *p*:0.703	0.908 (0.29–2.87); *p*:1.000
Female	217/302	5/3	58/57	63/60	1.00	0.431 (0.10–1.82); *p*:0.291	0.706 (0.47–1.06); *p*:0.091	0.684 (0.46–1.02); *p*:0.059	2.176 (0.52–9.18); *p*:0.305
Age									
< 63	387/527	4/8	75/97	79/105	1.00	1.469 (0.44–4.91); *p*:0.770	0.950 (0.68–1.32); *p*:0.759	0.976 (0.71–1.34); *p*:0.882	0.675 (0.20–2.26); *p*:0.573
≥ 63	452/573	6/2	101/103	107/105	1.00	0.263 (0.05–1.31); *p*:0.149	0.804 (0.60–1.09); *p*:0.156	0.774 (0.58–1.04); *p*:0.090	3.667 (0.74–18.24); *p*:0.151
Smoking status									
Never	463/811	6/7	105/142	111/149	1.00	0.666 (0.22–1.99); *p*:0.465	0.772 (0.59–1.02); *p*:0.067	0.766 (0.58–1.01); *p*:0.054	1.438 (0.48–4.30); *p*:0.513
Ever	376/289	4/3	71/58	75/61	1.00	0.976 (0.22–4.39); *p*:1.000	1.063 (0.73–1.55); *p*:0.753	1.058 (0.73–1.53); *p*:0.765	1.035 (0.23–4.66); *p*:1.000
Alcohol consumption									
Never	571/1024	7/9	125/184	132/193	1.00	0.717 (0.27–1.94); *p*:0.509	0.821 (0.64–1.05); *p*:0.120	0.815 (0.64–1.04); *p*:0.100	1.350 (0.50–3.64); *p*:0.552
Ever	268/76	3/1	51/16	54/17	1.00	1.175 (0.12–11.46); *p*:1.000	1.106 (0.60–2.05); *p*:0.748	1.110 (0.61–2.03); *p*:0.733	0.865 (0.09–8.42); *p*:1.000

[a] The genotyping success rate was 99.02% for rs3741172; [b] Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when n ≤ 5; Bonferroni correction was performed to correct the *p* value, bold values are statistically significant ($p < 0.0125$).

Linkage disequilibrium analyses and association test

Linkage disequilibrium analyses in both controls and cases were performed as shown in Tables 11, 12 and Figures 1, 2. There were strong correlations between these seven loci. Association test was conducted using Haploview software (version 4.2), strong associations were detected between these seven loci.

Haplotype analyses of ALDH3B2 polymorphisms and susceptibility of ESCC

As demonstrated in Table 13, haplotype analyses showed that ALDH3B2 $G_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}A_{rs9787887}$ was most common haplotype in both groups (66.3% in controls, 64.6% in cases). As compared with $G_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}A_{rs9787887}$, none of the haplotypes was associated with the susceptibility of ESCC.

Power calculation

The power calculation was performed using the "Power and Sample Size Calculation" software (http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize). Based on the assumption that the type I error probability for a two sided test α equals 0.05, the probability of exposure in controls P_0 is 0.126 in control. In the current study, using ligation detection reaction method, the successful rate of genotyping exceeded 98%. There were in total 1315 controls and 1043 cases successfully genotyped. The ratio of control/case (m) equals 1.261, and the correlation coefficient for exposure between matched case and controls (f) is 0.619. The power value is 0.936 as calculated by the "Power and Sample Size Calculation" software.

DISCUSSION

In this hospital-based case-control epidemiological study, we investigated the association between tagging SNPs of *ALDH3B2* and the risk of developing esophageal squamous cell carcinoma in a Chinese population. We found *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 polymorphisms were not implicated with altered susceptibility of ESCC according to age, gender, cigarette smoking and alcohol drinking stratification analyses.

Despite a suspected association between alcohol drinking and death to cancer reported in an epidemiological study as early as 1903, it took until 1988 for the research community to agree on the potential risk through the International Agency for Research on Cancer (IARC). Clear Patterns have emerged between alcohol consumption and esophageal cancer. Essentially, alcohol and its metabolite acetaldehyde were both designated as type 1A carcinogen [13]. The cytotoxic properties, the ability to form DNA-acetaldehyde adducts and to generate

Table 6: Stratified analyses between *ALDH3B2* rs4646823 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control) a				Adjusted OR b (95%CI); p				
	GG	TT	GT	TT+GT	GG	TT	GT	TT+GT	TT vs. (GG+GT)
Sex									
Male	10/12	598/762	137/174	735/936	1.00	1.062 (0.46–2.48); p:0.889	1.058 (0.44–2.52); p:0.898	1.061 (0.46–2.47); p:0.890	0.993 (0.78–1.26); p:0.954
Female	5/6	210/291	65/65	275/356	1.00	1.155 (0.35–3.83); p:1.000	0.833 (0.24–2.88); p:1.000	1.079 (0.33–3.57); p:1.000	0.732 (0.50–1.07); p:0.102
Age									
< 63	5/11	373/506	88/115	461/621	1.00	0.617 (0.21–1.79); p:0.450	0.594 (0.20–1.77); p:0.436	0.612 (0.21–1.77); p:0.450	0.999 (0.74–1.35); p:0.993
≥ 63	10/7	435/547	114/124	549/671	1.00	1.796 (0.68–4.76); p:0.232	1.554 (0.57–4.22); p:0.384	1.746 (0.66–4.62); p:0.255	0.840 (0.64–1.01); p:0.216
Smoking status									
Never	8/14	450/777	116/169	566/946	1.00	0.987 (0.41–2.37); p:0.976	0.833 (0.34–2.05); p:0.689	0.955 (0.40–2.29); p:0.918	0.855 (0.66–1.04); p:0.229
Ever	7/4	358/276	86/70	444/346	1.00	1.349 (0.39–4.66); p:0.764	1.424 (0.40–5.06); p:0.756	1.364 (0.40–4.70); p:0.764	1.032 (0.73–1.46); p:0.857
Alcohol consumption									
Never	10/17	556/978	137/222	693/1200	1.00	1.035 (0.47–2.28); p:0.932	0.953 (0.42–2.14); p:0.908	1.019 (0.46–2.24); p:0.963	0.924 (0.73–1.16); p:0.503
Ever	5/1	252/75	65/17	317/92	1.00	1.488 (0.17–12.94); p:1.000	1.308 (0.14–11.95); p:1.000	1.451 (0.17–12.58); p:1.000	0.864 (0.48–1.54); p:0.620

[a] The genotyping success rate was 99.02% for rs4646823; [b] Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n \leq 5$; Bonferroni correction was performed to correct the p value, bold values are statistically significant ($p < 0.0125$).

Table 7: Stratified analyses between *ALDH3B2* rs78402723 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control)a				Adjusted ORb (95%CI); p				
	GG	AA	AG	AA+AG	GG	AA	AG	AA+AG	AA vs. (GG+AG)
Sex									
Male	629/798	4/7	111/143	115/150	1.00	1.379 (0.40–4.73); p:0.764	0.985 (0.75–1.29); p:0.911	0.973 (0.75–1.27); p:0.837	0.727 (0.21–2.49); p:0.764
Female	220/301	5/3	55/58	60/61	1.00	0.439 (0.10–1.85); p:0.294	0.771 (0.51–1.16); p:0.210	0.743 (0.50–1.11); p:0.141	2.176 (0.52–9.18); p:0.305
Age									
< 63	392/526	4/8	70/98	74/106	1.00	1.490 (0.45–4.99); p:0.572	1.043 (0.75–1.46); p:0.803	1.068 (0.77–1.48); p:0.693	0.675 (0.20–2.26); p:0.573
≥ 63	457/573	5/2	96/103	101/105	1.00	0.319 (0.06–1.65); p:0.252	0.856 (0.63–1.16); p:0.315	0.829 (0.62–1.12); p:0.220	3.056 (0.59–15.81); p:0.255
Smoking status									
Never	467/810	6/7	101/143	107/150	1.00	0.673 (0.23–2.01); p:0.476	0.816 (0.62–1.08); p:0.154	0.808 (0.62–1.06); p:0.126	1.438 (0.48–4.30); p:0.513
Ever	382/289	3/3	65/58	68/61	1.00	1.322 (0.27–6.60); p:1.000	1.179 (0.80–1.73); p:0.401	1.186 (0.81–1.73); p:0.377	0.776 (0.16–3.87); p:1.000
Alcohol consumption									
Never	578/1023	7/9	118/185	125/194	1.00	0.726 (0.27–1.96); p:0.526	0.886 (0.69–1.14); p:0.346	0.877 (0.69–1.12); p:0.297	1.350 (0.50–3.64); p:0.552
Ever	271/76	2/1	48/16	50/17	1.00	1.783 (0.16–19.93); p:0.527	1.189 (0.64–2.21); p:0.585	1.212 (0.66–2.22); p:0.533	0.577 (0.05–6.43); p:0.535

[a]The genotyping success rate was 98.98% for rs78402723; [b]Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n \leq 5$; Bonferroni correction was performed to correct the p value, bold values are statistically significant ($p < 0.0125$).

additional mutagenic species at concentrations attainable *in vivo* may underlie the carcinogenic effects [22]. Most of the acetaldehyde generated during alcohol metabolism *in vivo* is rapidly eliminated by aldehyde dehydrogenase (ALDH), which renders ALDH an important role in carcinogen balancing and therefore carcinogenesis. In fact, there is ample evidence showed that subjects with an inactive form of ALDH2 (heterozygous for ALDH2 mutation) have an increased risk of developing various types of head and neck cancers as a consequent of intense exposure to acetaldehyde. Case-control studies of various Japanese drinking populations [23–29] and Chinese

Table 8: Stratified analyses between *ALDH3B2* rs7947978 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control) [a]				Adjusted OR[b] (95%CI); *p*				
	CC	AA	CA	AA+CA	CC	AA	CA	AA+CA	AA vs. (CC+CA)
Sex									
Male	595/762	7/12	143/174	150/186	1.00	1.339 (0.52–3.42); *p*:0.541	0.950 (0.74–1.22); *p*:0.683	0.968 (0.76–1.23); *p*:0.792	0.740 (0.29–1.89); *p*:0.527
Female	209/292	5/6	66/64	71/70	1.00	0.859 (0.26–2.85); *p*:1.000	0.694 (0.47–1.02); *p*:0.064	0.706 (0.49–1.03); *p*:0.068	1.079 (0.33–3.57); *p*:1.000
Age									
< 63	372/507	4/11	90/114	94/125	1.00	2.018 (0.64–6.39); *p*:0.295	0.929 (0.68–1.26); *p*:0.640	0.976 (0.72–1.32); *p*:0.872	0.489 (0.16–1.55); *p*:0.295
≥ 63	432/547	8/7	119/124	127/131	1.00	0.691 (0.25–1.92); *p*:0.476	0.823 (0.62–1.09); *p*:0.174	0.815 (0.62–1.07); *p*:0.143	1.392 (0.50–3.86); *p*:0.524
Smoking status									
Never	447/778	7/14	120/168	127/182	1.00	1.149 (0.46–2.87); *p*:0.766	0.824 (0.62–1.05); *p*:0.102	0.823 (0.64–1.06); *p*:0.134	0.834 (0.34–2.08); *p*:0.697
Ever	357/276	5/4	89/70	94/74	1.00	1.035 (0.28–3.89); *p*:1.000	1.017 (0.72–1.44); *p*:0.923	1.018 (0.72–1.44); *p*:0.918	0.970 (0.26–3.64); *p*:1.000
Alcohol consumption									
Never	552/979	9/17	142/221	151/238	1.00	1.065 (0.47–2.41); *p*:0.879	0.878 (0.69–1.11); *p*:0.276	0.889 (0.71–1.12); *p*:0.313	0.915 (0.41–2.07); *p*:0.831
Ever	252/75	3/1	67/17	70/18	1.00	1.120 (0.12–10.93); *p*:1.000	0.853 (0.47–1.54); *p*:0.597	0.864 (0.48–1.54); *p*:0.620	0.865 (0.09–8.42); *p*:1.000

[a] The genotyping success rate was 99.02% for rs7947978; [b] Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n \leq 5$; Bonferroni correction was performed to correct the *p* value, bold values are statistically significant ($p < 0.0125$).

Table 9: Stratified analyses between *ALDH3B2* rs866907 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control)[a]				Adjusted OR[b] (95% CI); *p*				
	AA	GG	AG	GG+AG	AA	GG	AG	GG+AG	GG vs. (AA+AG)
Sex									
Male	7/12	598/762	140/174	738/936	1.00	0.743 (0.29–1.90); *p*:0.534	0.725 (0.28–1.89); *p*:0.509	0.740 (0.29–1.89); *p*:0.527	0.993 (0.78–1.26); *p*:0.954
Female	5/6	211/291	64/65	275/356	1.00	1.149 (0.35–3.82); *p*:1.000	0.846 (0.25–2.91); *p*:1.000	1.079 (0.32–3.57); *p*:1.000	0.746 (0.51–1.09); *p*:0.126
Age									
< 63	4/11	373/506	89/115	462/621	1.00	0.493 (0.16–1.56); *p*:0.295	0.470 (0.15–1.53); *p*:0.281	0.489 (0.16–1.55); *p*:0.295	0.999 (0.74–1.35); *p*:0.993
≥ 63	8/7	436/547	115/124	551/671	1.00	1.434 (0.52–3.99); *p*:0.487	1.232 (0.43–3.51); *p*:0.695	1.392 (0.50–3.86); *p*:0.524	0.849 (0.64–1.12); *p*:0.245
Smoking status									
Never	7/14	451/777	116/169	567/946	1.00	0.861 (0.35–2.15); *p*:0.749	0.728 (0.29–1.86); *p*:0.506	0.834 (0.34–2.08); *p*:0.697	0.864 (0.67–1.12); *p*:0.262
Ever	5/4	358/276	88/70	446/346	1.00	0.964 (0.26–3.62); *p*:1.000	0.994 (0.26–3.84); *p*:1.000	0.970 (0.26–3.64); *p*:1.000	1.032 (0.73–1.46); *p*:0.857
Alcohol consumption									
Never	9/17	557/978	137/222	694/1200	1.00	0.930 (0.41–2.10); *p*:0.860	0.858 (0.37–1.98); *p*:0.719	0.915 (0.41–2.07); *p*:0.831	0.932 (0.74–1.17); *p*:0.551
Ever	3/1	252/75	67/17	319/92	1.00	0.893 (0.09–8.71); *p*:1.000	0.761 (0.07–7.78); *p*:1.000	0.865 (0.09–8.42); *p*:1.000	0.864 (0.48–1.54); *p*:0.620

[a] The genotyping success rate was 99.02% for rs866907; [b] Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n \leq 5$; Bonferroni correction was performed to correct the *p* value, bold values are statistically significant ($p < 0.0125$).

alcoholics [30] have consistently reported that the inactive ALDH2 encoded by the ALDH2*1/2*2 genotype is a strong risk factor for esophageal cancer. In resemblance with ALDH2, ALDH3 also plays a pivotal role in the alcohol metabolism, we thus hypothesize that ALDH3 family may be of potential relevance to carcinogenesis.

Table 10: Stratified analyses between *ALDH3B2* rs9787887 polymorphism and ESCC risk by sex, age, smoking status and alcohol consumption

Variable	(case/control) [a]				Adjusted OR [b] (95%CI); p				
	GG	AA	AG	AG+AA	GG	AA	AG	AG+AA	AA vs. (GG+AG)
Sex									
Male	84/113	319/420	338/414	657/834	1.00	0.979 (0.71–1.35); p:0.894	0.911 (0.66–1.25); p:0.562	0.944 (0.70–1.27); p:0.705	0.949 (0.78–1.15); p:0.593
Female	45/43	119/168	116/151	235/319	1.00	1.477 (0.92–2.39); p:0.110	1.362 (0.84–2.21); p:0.209	1.421 (0.91–2.23); p:0.126	0.854 (0.62–1.17); p:0.323
Age									
< 63	63/83	203/287	197/262	400/549	1.00	1.158 (0.80–1.67); p:0.433	1.090 (0.75–1.58); p:0.649	1.124 (0.80–1.59); p:0.506	0.939 (0.74–1.20); p:0.606
≥ 63	66/73	235/301	257/303	492/604	1.00	1.158 (0.80–1.68); p:0.442	1.066 (0.74–1.55); p:0.737	1.110 (0.78–1.58); p:0.563	0.909 (0.73–1.14); p:0.408
Smoking status									
Never	76/107	245/430	253/422	498/852	1.00	1.247 (0.89–1.74); p:0.194	1.241 (0.89–1.73); p:0.202	1.215 (0.89–1.66); p:0.223	0.916 (0.74–1.13); p:0.411
Ever	53/49	193/158	201/143	394/301	1.00	0.885 (0.57–1.38); p:0.589	0.770 (0.49–1.20); p:0.246	0.826 (0.55–1.25); p:0.369	0.923 (0.70–1.22); p:0.579
Alcohol consumption									
Never	91/141	305/552	306/523	611/1075	1.00	1.168 (0.87–1.57); p:0.307	1.103 (0.82–1.49); p:0.520	1.136 (0.86–1.51); p:0.376	0.924 (0.77–1.11); p:0.409
Ever	38/15	133/36	148/42	281/78	1.00	0.686 (0.34–1.38); p:0.291	0.719 (0.36–1.43); p:0.346	0.703 (0.37–1.35); p:0.285	1.132 (0.71–1.82); p:0.607

[a] The genotyping success rate was 98.81% for rs9787887; [b] Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model; Fisher exact test was performed when $n ≤ 5$; Bonferroni correction was performed to correct the p value, bold values are statistically significant ($p < 0.0125$).

Table 11: Linkage disequilibrium analyses of *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 in control group

D'	rs3741172	rs4646823	rs78402723	rs7947978	rs866907	rs9787887
rs34589365	**0.991**	**0.993**	**0.992**	**0.993**	**0.993**	**0.977**
rs3741172	-	1	1	**0.995**	**1**	**1**
rs4646823	-	-	1	**0.996**	**1**	**1**
rs78402723	-	-	-	**0.995**	**1**	**1**
rs7947978	-	-	-	-	**0.996**	**0.993**
rs866907	-	-	-	-	-	**1**

r^2	rs3741172	rs4646823	rs78402723	rs7947978	rs866907	rs9787887
rs34589365	0.013	0.017	0.013	0.017	0.017	0.272
rs3741172	-	**0.782**	**0.995**	**0.777**	**0.782**	0.182
rs4646823	-	-	**0.786**	**0.988**	**1**	0.233
rs78402723	-	-	-	**0.781**	**0.786**	0.183
rs7947978	-	-	-	-	**0.988**	0.229
rs866907	-	-	-	-	-	0.233

In line with our speculation, overexpression of ALDH3 protects cells from 4-hydroxynonenal induced apoptosis, suggesting a functional relevance of ALDH3 in the carcinogenesis. On the other hand, many of the sphere cells and stem cells reported in different organs have recently been found to be associated with elevated ALDH1A1 enzyme activity [31–33]. ALDH1A1 expression or activity may be used with other cell surface markers to identify tumor-initiating cells in hepatocellular, prostate and breast solid carcinomas [34–36]. ALDH1A1 has also been detected to be associated with early metastasis and poor clinical outcome [33]. In addition, the ALDH enzymes also play a pivotal role in epithelial homeostasis. Deregulation of these enzymes is associated with multiple cancers [37–

Table 12: Linkage disequilibrium analyses of *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 in case group

D'	rs3741172	rs4646823	rs78402723	rs7947978	rs866907	rs9787887
rs34589365	**0.984**	**0.982**	**0.983**	**0.979**	**0.981**	**0.969**
rs3741172	-	**1**	**1**	**1**	**1**	**0.989**
rs4646823	-	-	**1**	**0.985**	**1**	**0.957**
rs78402723	-	-	-	**1**	**1**	**1**
rs7947978	-	-	-	-	**1**	**0.974**
rs866907	-	-	-	-	-	**0.991**

r2	rs3741172	rs4646823	rs78402723	rs7947978	rs866907	rs9787887
rs34589365	0.014	0.017	0.013	0.017	0.017	0.248
rs3741172	-	**0.828**	**0.943**	**0.824**	**0.845**	0.194
rs4646823	-	-	**0.78**	**0.966**	**0.981**	0.219
rs78402723	-	-	-	**0.776**	**0.796**	0.186
rs7947978	-	-	-	-	**0.976**	0.228
rs866907	-	-	-	-	-	0.231

Table 13: *ALDH3B2* haplotype frequencies (%) in cases and controls and risk of ESCC

Haplotypes	Case (freq)	Control (freq)	Crude OR(95% CI)	*p*
*ALDH3B2*$A_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}G_{rs9787887}$	247 (12.1)	323 (12.4)	1.007 [0.840~1.206]	0.943
*ALDH3B2*$G_{rs34589365}A_{rs3741172}G_{rs4646823}A_{rs78402723}A_{rs7947978}A_{rs866907}G_{rs9787887}$	184 (9.0)	219 (8.4)	1.106 [0.898~1.363]	0.344
*ALDH3B2*$G_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}A_{rs9787887}$	1318 (64.6)	1735 (66.3)	1.000	1.000
*ALDH3B2*$G_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}G_{rs9787887}$	235 (11.5)	279 (10.6)	1.109 [0.919~1.338]	0.281

Haplotypes were composited by *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907, rs9787887. All those frequency < 3% were ignored in analysis, most common haplotype *ALDH3B2*$G_{rs34589365}G_{rs3741172}T_{rs4646823}G_{rs78402723}C_{rs7947978}G_{rs866907}A_{rs9787887}$ was selected as reference.

41]. In the current study, we demonstrated that there are no significant association between ALDH3B2 SNPs and risk of ESCC for the first time. More importantly, we have analyzed the interaction of the genetic background and environmental risk factors. These findings indicate that ALDH3B2 may not be the primary contributor to ESCC carcinogenesis. Considering the biological function of ALDH3B2, the evaluation with combined considerations based on both genetic and environmental factors would be much more precise and meaningful.

Cigarette smoking and alcohol drinking have emerged as well-known risk factors of ESCC. Although *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 were not associated with the susceptibility of ESCC in the current study, the cigarette smoking rate and alcohol drinking rate were significantly higher in the ESCC cases, exemplifying the significance of interaction between the environmental and genetic risk factors in causing esophageal squamous carcinoma.

Our finding that there are more male than female subjects in the case group was in consistent with the comprehensive data recently published by the National Office for Cancer Prevention and Control as well as the National Cancer Center of China. The fact that smoking and alcohol drinking are far more prevalent in male subjects implicated the importance of these risk factors in carcinogenesis of ESCC.

One of the limitations of our previous studies investigating the association between SNPs and risk of

ESCC was the sample size [12]. To overcome that, we have recruited a total of 2358 subjects including 1043 ESCC cases and 1315 controls in the current study. Yet, we do acknowledge there are still some limitations in this study. First, this study is limited by the sample sources, future studies designed and conduced in multiple ethnical populations and various geographic locations would be more convincing. Second, the lack of a validation cohort compromised the power of our study. Third, we are refrained by the lack of technical support to establish single nucleotide mutation cell/animal model, the actual biological function of these SNPs in esophageal carcinoma remains obscure, the underlying mechanisms are yet to be further dissected. Last but not least, the detailed information with regard to cancer metastasis and survival were not provided, the correlations between SNPs and outcomes have not been analyzed as this follow-up is still ongoing.

MATERIALS AND METHODS

Ethics statement

This hospital-based case-control study was approved by the Review Board of Jiangsu University (Zhenjiang, China). All subjects provided written informed consents. This study has complied with the World Medical Association Declaration of Helsinki with regard to ethical
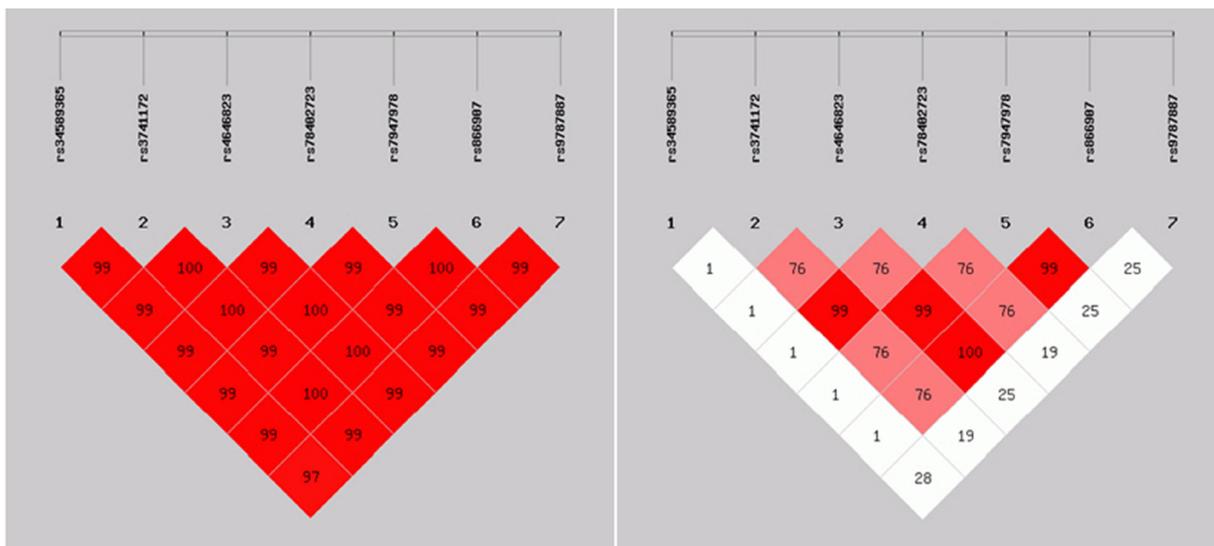
Figure 1: $D' > 0$, $r^2 > 0$: There were linkage disequilibrium correlations among different loci; $D' > 70\%$, $r^2 > 30\%$. There were closer linkage disequilibrium correlations among different loci.
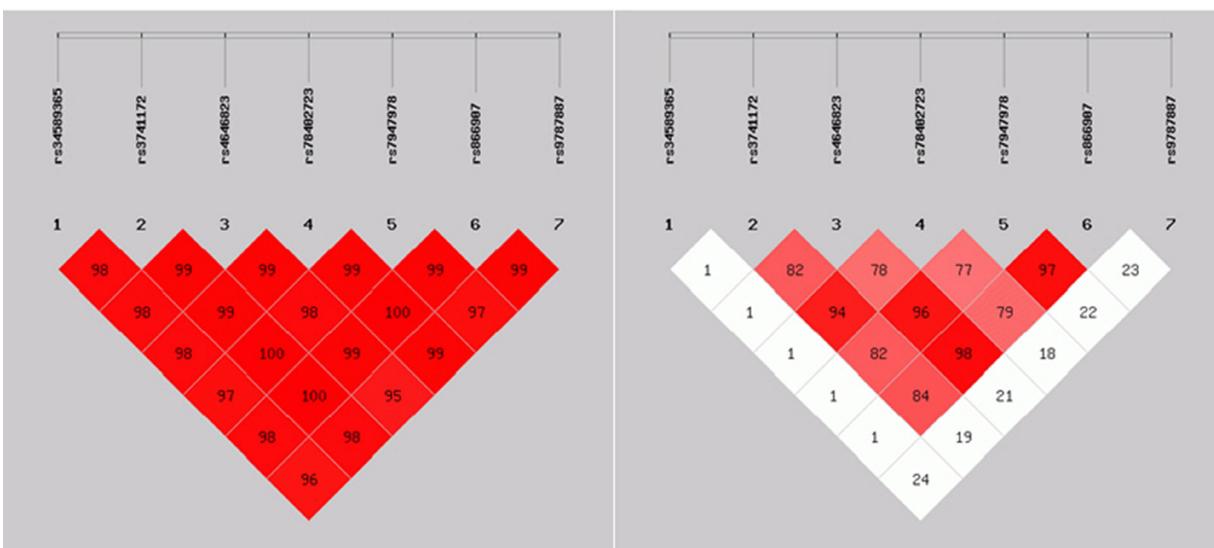
Figure 2: $D' > 0$, $r^2 > 0$: There were linkage disequilibrium correlations among different loci; $D' > 70\%$ (0.7), $r^2 > 30\%$ (0.3). There were closer linkage disequilibrium correlations among different loci.

conduct of research involving human subjects and/or animals.

Study populations

A total of subjects consisting of 1043 ESCC cases and 1315 non-cancer controls frequency-matched to the cases regarding age and gender (p = 0.121 and 0.880, respectively) were enrolled in this study (Table 1). All patients and controls were consecutively recruited from the Affiliated People's Hospital of Jiangsu University (Zhenjiang, China) from October 2008 to January 2017. All cases of esophageal cancer were diagnosed as ESCC histologically. The exclusion criteria included cancer history, metastasized cancer or chemotherapy/radiotherapy history.

Each subject was individually questioned by experienced interviewers with a questionnaire to obtain information on demographic information and related risk factors (including alcohol consumption and cigarette smoking). After written informed consent was provided, two milliliters of venous blood were collected from each subject. The "Smokers" cohort included individuals who smoked one cigarette per day for more than one year. Subjects who had more than three alcoholic drinks a week for more than six months were included in the "Alcohol drinkers" cohort.

Genomic DNA extraction, SNP selection and genotyping

Genomic DNA was isolated from peripheral blood using QIAamp DNA Blood Mini Kit (Qiagen, Berlin, Germany) as reported [12]. Sample DNA were amplified by PCR according to the manufacturer's protocol. Gene polymorphisms were analyzed by the ligation detection reaction (LDR) method with technical support from Genesky Biotechnology Inc. (Shanghai, China). 10% of the total samples were randomly selected for repeated analyses for quality control. Pilot linkage disequilibrium analyses were performed in the Chinese Han population to choose the SNP loci with moderate correlation, and tag SNPs were selected for further analyses.

Statistical analyses

Statistical analyses were conducted using SPSS 23.0 statistical package (SPCC Inc., Chicago, IL). Hardy-Weinberg equilibrium for genotypes was tested by goodness-of-fit χ^2 in control group. Variations of demographic characteristics and genotypes of the *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 between the controls and cases were evaluated using the *chi*-square (χ^2) test to examine the statistical differences. The associations between these seven SNPs and risk

of ESCC were analyzed by PLINK software (v1.07, available at http://zzz.bwh.harvard.edu/plink/download.shtml). Crude ORs and adjusted ORs when adjusting for age, sex, smoking and alcohol drinking status were also computed using logistic regression analyses. Bilateral probability tests were taken, p value < 0.05 was considered statistically significant.

CONCLUSIONS

The esophageal squamous carcinoma is associated with various factors including gene, environment and life-style. Our findings that *ALDH3B2* rs34589365, rs3741172, rs4646823, rs78402723, rs7947978, rs866907 and rs9787887 polymorphisms were not implicated with altered susceptibility of ESCC in different age, gender, cigarette smoking and alcohol drinking status, when interpreted with caution, could be helpful in evaluating the susceptibility to ESCC.

REFERENCES

1. Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, Allen C, Hansen G, Woodbrook R, Wolfe C, Hamadeh RR, Moore A, Werdecker A, et al;

Global Burden of Disease Cancer Collaboration. The Global Burden of Cancer 2013. JAMA Oncol. 2015; 1:505–27. https://doi.org/10.1001/jamaoncol.2015.0735.

2. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet. 2017; 390:2383–2396. https://doi.org/10.1016/S0140-6736(17)31462-9.

3. Xu Y, Yu X, Chen Q, Mao W. Neoadjuvant versus adjuvant treatment: which one is better for resectable esophageal squamous cell carcinoma? World J Surg Oncol. 2012; 10:173. https://doi.org/10.1186/1477-7819-10-173.

4. Gammon MD, Schoenberg JB, Ahsan H, Risch HA, Vaughan TL, Chow WH, Rotterdam H, West AB, Dubrow R, Stanford JL, Mayne ST, Farrow DC, Niwa S, et al. Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. J Natl Cancer Inst. 1997; 89:1277–84.

5. Lee CH, Wu DC, Lee JM, Wu IC, Goan YG, Kao EL, Huang HL, Chan TF, Chou SH, Chou YP, Lee CY, Chen PS, Ho CK, et al. Carcinogenetic impact of alcohol intake on squamous cell carcinoma risk of the oesophagus in relation to tobacco smoking. Eur J Cancer. 2007; 43:1188–99. https://doi.org/10.1016/j.ejca.2007.01.039.

6. De Stefani E, Barrios E, Fierro L. Black (air-cured) and blond (flue-cured) tobacco and cancer risk. III: Oesophageal cancer. Eur J Cancer. 1993; 29A:763–6.

7. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med. 2003; 349:2241–52. https://doi.org/10.1056/NEJMra035010.

8. Brown LM, Hoover R, Silverman D, Baris D, Hayes R, Swanson GM, Schoenberg J, Greenberg R, Liff J, Schwartz A, Dosemeci M, Pottern L, Fraumeni JF Jr. Excess incidence of squamous cell esophageal cancer among US Black men: role of social class and other risk factors. Am J Epidemiol. 2001; 153:114–22.

9. Taylor PR, Qiao YL, Abnet CC, Dawsey SM, Yang CS, Gunter EW, Wang W, Blot WJ, Dong ZW, Mark SD. Prospective study of serum vitamin E levels and esophageal and gastric cancers. J Natl Cancer Inst. 2003; 95:1414–6.

10. Abnet CC, Lai B, Qiao YL, Vogt S, Luo XM, Taylor PR, Dong ZW, Mark SD, Dawsey SM. Zinc concentration in esophageal biopsy specimens measured by x-ray fluorescence and esophageal cancer risk. J Natl Cancer Inst. 2005; 97:301–6. https://doi.org/10.1093/jnci/dji042.

11. Matejcic M, Gunter MJ, Ferrari P. Alcohol Metabolism and Oesophageal Cancer: A Systematic Review of the Evidence. Carcinogenesis. 2017; 38:859–872. https://doi.org/10.1093/carcin/bgx067.

12. Yin J, Wang L, Tang W, Wang X, Lv L, Shao A, Shi Y, Ding G, Chen S, Gu H. RANK rs1805034 T>C polymorphism is associated with susceptibility of esophageal cancer in a Chinese population. PLoS One. 2014; 9:e101705. https://doi.org/10.1371/journal.pone.0101705.

13. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Alcohol consumption and ethyl carbamate. IARC Monogr Eval Carcinog Risks Hum. 2010; 96:3–1383.

14. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012; 100:1–538.

15. Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V; WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. Lancet Oncol. 2009; 10:1033–4.

16. Zakhari S. Overview: how is alcohol metabolized by the body? Alcohol Res Health. 2006; 29:245–54.

17. Ji Y, Li X, Li Y, Zhong Y, Cao J, Xu R, Wang J, Zhou F, Li X, Yu D, Wen JG, Goscinski MA, Zhang M, et al. Aldehyde Dehydrogenase-1 Expression Predicts Unfavorable Outcomes in Patients with Esophageal Squamous Cell Carcinoma. Anticancer Res. 2016; 36:343–9.

18. Crabb DW, Edenberg HJ, Bosron WF, Li TK. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. J Clin Invest. 1989; 83:314–6. https://doi.org/10.1172/JCI113875.

19. Liu K, Song G, Zhu X, Yang X, Shen Y, Wang W, Shi G, Li Q, Duan Y, Zhao Y, Feng G. Association between ALDH2 Glu487Lys polymorphism and the risk of esophageal cancer. Medicine (Baltimore). 2017; 96:e6111. https://doi.org/10.1097/MD.0000000000006111.

20. Yukawa Y, Muto M, Hori K, Nagayoshi H, Yokoyama A, Chiba T, Matsuda T. Combination of ADH1B*2/ALDH2*2 polymorphisms alters acetaldehyde-derived DNA damage in the blood of Japanese alcoholics. Cancer Sci. 2012; 103:1651–5. https://doi.org/10.1111/j.1349-7006.2012.02360.x.

21. Ma WJ, Lv GD, Zheng ST, Huang CG, Liu Q, Wang X, Lin RY, Sheyhidin I, Lu XM. DNA polymorphism and risk of esophageal squamous cell carcinoma in a population of North Xinjiang, China. World J Gastroenterol. 2010; 16:641–7.

22. Marietta C, Thompson LH, Lamerdin JE, Brooks PJ. Acetaldehyde stimulates FANCD2 monoubiquitination, H2AX phosphorylation, and BRCA1 phosphorylation in human cells in vitro: implications for alcohol-related carcinogenesis. Mutat Res. 2009; 664:77–83. https://doi.org/10.1016/j.mrfmmm.2009.03.011.

23. Yokoyama A, Muramatsu T, Ohmori T, Higuchi S, Hayashida M, Ishii H. Esophageal cancer and aldehyde dehydrogenase-2 genotypes in Japanese males. Cancer Epidemiol Biomarkers Prev. 1996; 5:99–102.

24. Hori H, Kawano T, Endo M, Yuasa Y. Genetic polymorphisms of tobacco- and alcohol-related metabolizing enzymes and human esophageal squamous cell carcinoma susceptibility. J Clin Gastroenterol. 1997; 25:568–75.

25. Yokoyama A, Muramatsu T, Ohmori T, Yokoyama T, Okuyama K, Takahashi H, Hasegawa Y, Higuchi S,

Maruyama K, Shirakura K, Ishii H. Alcohol-related cancers and aldehyde dehydrogenase-2 in Japanese alcoholics. Carcinogenesis. 1998; 19:1383–7.

26. Tanabe H, Ohhira M, Ohtsubo T, Watari J, Yokota K, Kohgo Y. Genetic polymorphism of aldehyde dehydrogenase 2 in patients with upper aerodigestive tract cancer. Alcohol Clin Exp Res. 1999; 23:17S–20S.

27. Yokoyama A, Muramatsu T, Omori T, Yokoyama T, Matsushita S, Higuchi S, Maruyama K, Ishii H. Alcohol and aldehyde dehydrogenase gene polymorphisms and oropharyngolaryngeal, esophageal and stomach cancers in Japanese alcoholics. Carcinogenesis. 2001; 22:433–9.

28. Matsuo K, Hamajima N, Shinoda M, Hatooka S, Inoue M, Takezaki T, Tajima K. Gene-environment interaction between an aldehyde dehydrogenase-2 (ALDH2) polymorphism and alcohol consumption for the risk of esophageal cancer. Carcinogenesis. 2001; 22:913–6.

29. Itoga S, Nomura F, Makino Y, Tomonaga T, Shimada H, Ochiai T, Iizasa T, Baba M, Fujisawa T, Harada S. Tandem repeat polymorphism of the CYP2E1 gene: an association study with esophageal cancer and lung cancer. Alcohol Clin Exp Res. 2002; 26:15S-9S. https://doi.org/10.1097/01.ALC.0000026828.13868.B5.

30. Chao YC, Wang LS, Hsieh TY, Chu CW, Chang FY, Chu HC. Chinese alcoholic patients with esophageal cancer are genetically different from alcoholics with acute pancreatitis and liver cirrhosis. Am J Gastroenterol. 2000; 95:2958–64. https://doi.org/10.1111/j.1572-0241.2000.02328.x.

31. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur MH, Diebel ME, Monville F, Dutcher J, Brown M, Viens P, Xerri L, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer Res. 2009; 69:1302–13. https://doi.org/10.1158/0008-5472.CAN-08-2741.

32. Carpentino JE, Hynes MJ, Appelman HD, Zheng T, Steindler DA, Scott EW, Huang EH. Aldehyde dehydrogenase-expressing colon stem cells contribute to tumorigenesis in the transition from colitis to cancer. Cancer Res. 2009; 69:8208–15. https://doi.org/10.1158/0008-5472.CAN-09-1132.

33. Charafe-Jauffret E, Ginestier C, Iovino F, Tarpin C, Diebel M, Esterni B, Houvenaeghel G, Extra JM, Bertucci F, Jacquemier J, Xerri L, Dontu G, Stassi G, et al. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. Clin Cancer Res. 2010; 16:45–55. https://doi.org/10.1158/1078-0432.CCR-09-1630.

34. Colombo F, Baldan F, Mazzucchelli S, Martin-Padura I, Marighetti P, Cattaneo A, Foglieni B, Spreafico M, Guerneri S, Baccarin M, Bertolini F, Rossi G, Mazzaferro V, et al. Evidence of distinct tumour-propagating cell populations with different properties in primary human hepatocellular carcinoma. PLoS One. 2011; 6:e21369. https://doi.org/10.1371/journal.pone.0021369.

35. van den Hoogen C, van der Horst G, Cheung H, Buijs JT, Lippitt JM, Guzman-Ramirez N, Hamdy FC, Eaton CL, Thalmann GN, Cecchini MG, Pelger RC, van der Pluijm G. High aldehyde dehydrogenase activity identifies tumor-initiating and metastasis-initiating cells in human prostate cancer. Cancer Res. 2010; 70:5163–73. https://doi.org/10.1158/0008-5472.CAN-09-3806.

36. Buijs JT, van der Horst G, van den Hoogen C, Cheung H, de Rooij B, Kroon J, Petersen M, van Overveld PG, Pelger RC, van der Pluijm G. The BMP2/7 heterodimer inhibits the human breast cancer stem cell subpopulation and bone metastases formation. Oncogene. 2012; 31:2164–74. https://doi.org/10.1038/onc.2011.400.

37. van den Hoogen C, van der Horst G, Cheung H, Buijs JT, Pelger RC, van der Pluijm G. The aldehyde dehydrogenase enzyme 7A1 is functionally involved in prostate cancer bone metastasis. Clin Exp Metastasis. 2011; 28:615–25. https://doi.org/10.1007/s10585-011-9395-7.

38. Marcato P, Dean CA, Pan D, Araslanova R, Gillis M, Joshi M, Helyer L, Pan L, Leidal A, Gujar S, Giacomantonio CA, Lee PW. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells. 2011; 29:32–45. https://doi.org/10.1002/stem.563.

39. Marchitti SA, Orlicky DJ, Brocker C, Vasiliou V. Aldehyde dehydrogenase 3B1 (ALDH3B1): immunohistochemical tissue distribution and cellular-specific localization in normal and cancerous human tissues. J Histochem Cytochem. 2010; 58:765–83. https://doi.org/10.1369/jhc.2009.95577310.1369/jhc.2010.955773.

40. Marcato P, Dean CA, Giacomantonio CA, Lee PW. Aldehyde dehydrogenase: its role as a cancer stem cell marker comes down to the specific isoform. Cell Cycle. 2011; 10:1378–84. https://doi.org/10.4161/cc.10.9.15486.

41. Zhang Q, Taguchi A, Schliekelman M, Wong CH, Chin A, Kuick R, Misek DE, Hanash S. Comprehensive proteomic profiling of aldehyde dehydrogenases in lung adenocarcinoma cell lines. Int J Proteomics. 2011; 2011:145010. https://doi.org/10.1155/2011/145010.