Pharmacogenomics and Personalized Medicine

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

The Relationship Between Single Nucleotide Polymorphisms of *SMAD3/SMAD6* and Risk of Esophageal Squamous Cell Carcinoma in Chinese Population

This article was published in the following Dove Press journal:
Pharmacogenomics and Personalized Medicine

Jinjie Yu[1,*]
Yunpeng Dong[1,*]
Weifeng Tang[2]
Huiwen Pan[2]
Lu Lv[2]
Tao Long[2]
Qiang Zhou[2]
Junqing Qi[2]
Jianchao Liu[2]
Guowen Ding[2]
Jun Yin[1]
Lijie Tan[1]

[1]Department of Thoracic Surgery, Zhongshan Hospital of Fudan University, Shanghai, People's Republic of China; [2]Department of Cardiothoracic Surgery, Affiliated People's Hospital of Jiangsu University, Jiangsu, People's Republic of China

*These authors contributed equally to this work

Background: The TGF-β signal pathways play a key role in the development and promotion of squamous cell carcinoma (SCC). The pathway is mediated by the *SMAD* family proteins that include *SMAD3* and *SMAD6*. Our study aimed to evaluate the relationship between single nucleotide polymorphism (SNP) of *SMAD3/SMAD6* and susceptibility to esophageal squamous cell carcinoma (ESCC) in the Chinese population.

Patients and Methods: This was a hospital-based case–control study compromised of 1043 ESCC patients and 1315 non-cancer patients. Seven *SMAD3/SMAD6* (rs8028147, rs3743343, rs3743342, rs8025774, rs8031440, rs803167, and rs34643453) SNPs were selected and used to evaluate their correlation with ESCC susceptibility. Genetic model tests, stratified analyses, linkage disequilibrium analyses, and haplotype analyses were performed in our study.

Results: Participants with *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A or rs8031627 G>A had a significantly higher risk of ESCC. This was more evident in males, older patients (>63 years), smokers, and non-alcohol drinking participants. Linkage disequilibrium analyses further revealed that there were strong correlations between *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A, and rs8031627 G>A. In the same line, haplotype analyses revealed that *SMAD3* ACCCGG*SMAD6*A and *SMAD3*AGCCGG*SMAD6*A were associated with less susceptibility to ESCC while *SMAD3*ATTTAA*SMAD6*A was associated with a higher risk of ESCC.

Conclusion: SNPs of *SMAD3* were related to higher susceptibility to ESCC. As such, they may contribute to the development of viable strategies for early diagnosis and treatment of ESCC. However, more detailed association mechanisms between *SMAD3/SMAD6* SNPs and ESCC need further experiments to prove.

Keywords: *SMAD*, single nucleotide polymorphism, esophageal squamous cell carcinoma

Introduction

Small mothers against decapentaplegic (*SMAD*) proteins are a family of transducers and transcriptional modulators that functions by transforming the growth factor beta (TGF-β) signaling pathway. Alterations of the TGF-β signal pathway lead to uncontrolled cell proliferation, thereby contributing to oncogenesis.[1] Ligands of the TGF-β family bind to complexes formed by its type I and type II receptors to trigger the pathway. The activated type I receptors then phosphorylate receptor-activated *SMAD*s (R-*SMAD*s, including *SMAD*2, −3, −5, −8) which then interact with *SMAD*4

Correspondence: Jun Yin; Lijie Tan
Department of Thoracic Surgery, Zhongshan Hospital of Fudan University, No. 180, Fenglin Road, Xuhui District, Shanghai, People's Republic of China
Tel +86 13917483128; +86 13681972151
Email yin.jun2@zs-hospital.sh.cn; tan.lijie@zs-hospital.sh.cn

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/PGPM.S250076

(also known as the co-*SMAD*) to form a heterotrimeric complex that consists of two R-SMADs and one co-*SMAD*.[1,2] A subgroup of inhibitory *SMAD*s (I-*SMAD*s, including *SMAD*6, −7) attenuates the signaling by competing with the R-SMADs for TGF-β receptors or directly interacting with R-SMADs.[1,3,4]

Drugs that mediate the TGF-β signaling pathway have shown increasing potency of changing the current treatment of various cancers.[5] Given that the *SMAD*s family play a key role in activation or inhibition of the TGF-β signaling pathway, further studies on SMADs are becoming essential and promising.[6] *SMAD3* has been proven to have diverse correlations with different types of malignant tumors. For example, *SMAD3* impairment is associated with dysregulated cell proliferation and apoptosis in ovarian granulosa cell tumors.[7] In human colorectal cancer, linker phosphorylation of *SMAD3* is associated with tumor metastasis.[8] Inhibition of *SMAD3* diminishes the invasion and metastatic ability of breast cancer.[9,10] In addition, the expression of *SMAD3* suppressed tumor

Table 1 Distribution of Selected Demographic Variables and Risk Factors in ESCC Case and Control Groups

	Case Group (n=1043)		**Control Group (n=1315)**		***P***
Age (years) Mean ± SD	63.07 ± 7.27		62.88 ± 9.74		0.607
Age (years) <63	471	45.16%	636	48.37%	0.121
>63	572	54.84%	679	51.63%	
Gender Male	758	72.67%	952	72.40%	0.88
Female	285	27.33%	343	26.08%	
Smoking status Never	589	56.47%	964	73.31%	<0.001
Ever	454	43.53%	351	26.69%	
Alcohol consumption Never	714	68.46%	1222	92.93%	<0.001
Ever	329	31.54%	93	7.07%	

Table 2 Primary Information of the Selected SNPs

Genotyped SNP	**rs8028147**	**rs3743343**	**rs3743342**	**rs8025774**	**rs8031440**	**rs8031627**	**rs34643453**
Ancestral Allele	**G**	**T**	**C**	**C**	**G**	**G**	**G**
Gene	*SMAD3*(4088)						*SMAD6*(4091)
Function	utr variant 5 prime	utr variant 3 prime					utr variant 5 prime
Regulome DB Score[a]	4	5	3a	4	4	4	4
TFBS[b]	–	–	–	–	–	–	Y
nsSNP	–	–	–	–	–	–	–
Chr Pos	67,125,895	67,194,437	67,193,329	67,190,938	67,191,641	67,191,781	66,702,737
Chromosome	15	15	15	15	15	15	15x`
MAF for Chinese in database(HAPMAP)	G = 0.440	C = 0.314	T = 0.453	T = 0.451	A = 0.463	A = 0.451	A = 0.268
MAF in the Controls	G=0.416	C=0.317	T=0.462	T=0.456	A=0.461	A=0.455	A=0.255
P value for HWE test in the controls	0.8066	0.1688	0.7432	0.7786	0.7641	0.8574	0.7750
Genotyping method	LDR						
%Genotyping value	99.02%	99.02%	99.02%	99.02%	98.98%	99.02%	95.59%

Notes: [a]http://www.regulomedb.org/; [b]TFBS, transcription factor binding site (http://snpinfo.niehs.nih.gov).

Table 3 Main Effects of Rs8028147 G>A, Rs3743343 T>C, Rs3743342 C>T, Rs8025774 C>T, Rs8031440 G>A, Rs8031627 G>A, Rs3463453 G>A on Risk of ESCC

Locus	Genotype	Control	Case	Co-Dominant		Dominant	Recessive	Allelic Test
				OR (95% CI)	*P*	OR (95% CI)/*P*	OR (95% CI)/*P*	OR (95% CI)/*P*
rs8028147	GG GA AA	225 641 444	169 499 357	1 1.070 (0.839–1.365) 1.036(0.822–1.306)	0.854 0.583 0.762	1.050(0.844–1.307) 0.660	1.042(0.878–1.238) 0.636	1.034(0.920–1.163) 0.576
rs3743343	TT TC CC	600 589 121	498 436 91	1 0.515(0.673–1.219) 0.191(0.751–1.059)	0.407 0.515 0.191	0.894(0.759–1.053) 0.181	0.957(0.720–1.273) 0.765	0.929(0.820–1.053) 0.249
rs3743342	CC TC TT	377 657 276	243 509 273	1 1.202(0.986–1.466) 1.535(1.216–1.936)	0.001 0.069 <0.001	1.300(1.078–1.568) 0.006	1.360(1.123–1.648) 0.002	1.237(1.102–1.389) <0.001
rs8025774	CC TC TT	385 655 270	253 503 269	1 1.169(0.960–1.423) 1.516(1.203–1.911)	0.002 0.120 <0.001	1.270(1.055–1.528) 0.011	1.371(1.130–1.662) 0.001	1.230(1.096–1.381) <0.001
rs8031440	GG GA AA	377 656 276	250 499 276	1 1.147(0.941–1.398) 1.508(1.197–1.900)	0.002 0.174 <0.001	1.254(1.041–1.510) 0.017	1.379(1.139–1.670) 0.001	1.228(1.094–1.379) 0.001
rs8031627	GG GA AA	387 653 270	252 502 271	1 1.181(0.970–1.437) 1.541(1.223–1.943)	0.001 0.098 <0.001	1.286(1.069–1.548) 0.008	1.384(1.142–1.679) 0.001	1.241(1.106–1.394) <0.001
rs3464453	GG GA AA	681 471 78	578 376 70	1 0.937(0.787–1.117) 1.056(0.751–1.485)	0.687 0.468 0.755	0.954(0.807–1.127) 0.581	1.084(0.776–1.513) 0.637	0.983(0.859–1.124) 0.798

development in gastric cancer.[11] Similarly, *SMAD6* has been demonstrated to be related to poor prognosis in non-small cell lung cancer.[12,13] However, there is only little evidence that shows *SMAD3* or *SMAD6* contribute to the progression of ESCC, especially the relationship between *SMAD* gene and ESCC. Cognizant to this, our study aimed to explore whether SNPs in SMAD3 or *SMAD6* had an effect on the risk of ESCC.

Patients and Methods

Ethical Statement

This was a case–control study approved by the Ethics Review Board of Jiangsu University (Zhenjiang, China). All study participants gave written informed consent prior to the study. The study also complied with the World Medical Association Declaration of Helsinki regarding ethical aspects of research involving human/animal subjects.

Study Population

A total of 2358 participants were enrolled in the study, of which 1043 were ESCC cases and 1315 non-cancer individuals acted as the controls. Frequency-matching of both groups was done based on gender and age. All participants were consecutively recruited from the Affiliated People's Hospital of Jiangsu University and Affiliated Hospital of Jiangsu University (Zhenjiang, China) between October 2008 and January 2017. All cases of ESCC were diagnosed histologically. Patients with a history of malignant tumor, metastasized cancer, and/or those who had undergone chemotherapy/radiotherapy were excluded from the study.

Individual interviews were conducted using a questionnaire well formulated to collect data of all the necessary demographic information and related risk factors. Then, participants gave 2 mL venous blood samples for subsequent tests. Participants who smoked at least one cigarette per day for the last one year or more were categorized as the "tobacco consumption" subgroup while those who drank more than three alcoholic drinks per week for the past six months or more were categorized as the "alcohol consumption" subgroup.

DNA Extraction and SNP Genotyping

Genomic DNA was isolated from the collected venous blood samples using the QIAamp DNA Blood Mini Kit

(Qiagen, Berlin, Germany) following the manufacturer's protocol.[14] The genotypes of 7 SNPs were then analyzed using the ligation detection reaction (LDR) method with technical support from Genesky Biotechnology Inc. (Shanghai, China). Quality control was done by repeating the analyses using 10% of randomly selected samples. 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 Analysis

All statistical analyses were conducted using the SPSS 25.0 statistical software package (SPCC Inc., Chicago, IL). Hardy–Weinberg equilibrium (HWE) for the genotypes was tested by goodness-of-fit χ^2 in the control group. Variations in demographic characteristics and genotypes of the *SMAD3* rs8028147, rs3743343, rs3743342, rs8025774, rs8031440, rs8031627, and *SMAD6* rs34643453 between both groups were evaluated using the χ^2 test to determine their statistical differences. The associations between these 7 SNPs and risk of ESCC were assessed by odds ratio (OR) and 95% confidence interval (CI) using logistics regression analyses for crude ORs and adjusted ORs according to age, sex, and tobacco and alcohol consumption status. Two-sided P value <0.05 was considered as statistically significant.

Results

Characteristics of the Study Population

The demographic information and risk factors of the participants are shown in Table 1. Age and gender were well matched in cases and control groups. However, the smoking and alcohol drinking statuses were significantly different between the groups (P<0.01).

Table 2 provides primary information of the 7 genotyped SNPs. The success rates of SNP genotyping were variable and the rates ranged between 95% and 100%. The minor allele frequency (MAF) of SNPs in the control group corresponded with that of the Chinese Han population provided by HapMap. Deviation tests for the HWE revealed that the control population was in the Hardy–

Table 4 Stratified Analyses Between Rs3743342 Polymorphism and ESCC Risk by Gender, Age, Smoking Status, and Alcohol Consumption

Variables		Control/Case				Adjusted OR/*P* (95% CI of OR)				
		CC	TC	TT	TC+TT	CC	TC	TT	TC+TT	TT vs (TC+CC)
Gender										
	Male	283/169	476/370	189/206	665/576	1.00	1.235/0.105 (0.957–1.592)	1.838/<0.001 (1.366–2.474)	1.404/0.006 (1.105–1.784)	0.624/<0.001 (0.489–0.797)
	Female	94/74	181/139	87/67	268/206	1.00	0.967/0.863 (0.661–1.414)	0.970/0.892 (0.622–1.513)	0.968/0.858 (0.677–1.384)	1.009/0.962 (0.697–1.460)
Age										
	<63	195/123	298/231	139/112	437/343	1.00	1.276/0.118 (0.940–1.732)	1.373/0.085 (0.957–1.971)	1.306/0.068 (0.980–1.741)	0.849/0.292 (0.626–1.151)
	≥63	182/120	359/278	137/161	496/439	1.00	1.049/0.751 (0.783–1.405)	1.655/0.004 (1.179–2.323)	1.215/0.167 (0.922–1.602)	0.624/0.001 (0.474–0.822)
Smoking Status										
	Never	277/140	473/282	210/152	683/434	1.00	1.173/0.218 (0.910–1.513)	1.438/0.016 (1.070–1.931)	1.254/0.063 (0.987–1.593)	0.772/0.037 (0.605–0.984)
	Ever	100/103	184/227	66/121	250/348	1.00	1.067/0.735 (0.731–1.559)	1.731/0.018 (1.100–2.275)	1.237/0.244 (0.865–1.771)	0.603/0.009 (0.414–0.879)
Alcohol consumption										
	Never	351/174	608/339	258/190	866/529	1.00	1.113/0.356 (0.887–1.396)	1.472/0.004 (1.132–1.914)	1.220/0.068 (0.986–1.509)	0.728/0.004 (0.586–0.905)
	Ever	26/69	49/170	18/83	67/253	1.00	1.279/0.397 (0.723–2.264)	1.723/0.126 (0.859–3.456)	1.400/0.224 (0.814–2.410)	0.687/0.205 (0.385–1.228)

Weinberg proportions for all the 7 SNPs with a significance level of 0.05.

The Risk of ESCC Associated with SNPs

The association between the risk of ESCC and each SNP is presented in Table 3. It presents the results obtained from the analysis of the association between the risk of ESCC and each SNP. The co-dominant model test, dominant model test, recessive model test, and an allelic test revealed that among the selected SNPs, *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A, and rs8031627 G>A were closely associated with a higher risk of ESCC in the case groups ($P<0.05$).

Stratification Analyses of *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A, rs8031627 G>A and the Risk of ESCC

Stratification analyses were conducted to further access *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A, and rs8031627 G>A on the risk of ESCC in the different subgroups based on gender, age, and smoking and alcohol drinking status.

The results of stratification analyses share a remarkable similarity (Tables 4–7). In males, participants older than 63 years, smokers, and non-alcohol drinking participants, almost all the mutant homozygotes of these four SNPs (except rs8031440 G>A in smokers) had a significantly higher likelihood of having ESCC ($P<0.05$).

Linkage Disequilibrium Analyses and Association Test

Linkage disequilibrium analyses of both the control and case groups are set out in supplementary Tables 1 and 2. There were significant correlations between *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440 G>A and rs8031627 G>A and risk of ESCC.

Haplotype Analysis of Polymorphisms and Susceptibility to ESCC

As summarized in Table 8, haplotype analysis of 7 SNPs showed that *SMAD3*$A_{rs8028147}C_{rs3743343}C_{rs3743342}C_{rs8025774}$

Table 5 Stratified Analyses Between Rs8025774 Polymorphism and ESCC Risk by Gender, Age, Smoking Status, and Alcohol Consumption

Variables		Control/Case				Adjusted OR/P (95% CI of OR)				
		CC	TC	TT	TC+TT	CC	TC	TT	TC+TT	TT vs (TC+CC)
Gender										
	Male	291/174	474/367	183/204	657/571	1.00	1.224/0.117 (0.951–1.575)	1.854/<0.001 (1.379–2.493)	1.397/0.006 (1.102–1.772)	0.615/<0.001 (0.480–0.786)
	Female	94/79	181/136	87/65	268/201	1.00	0.888/0.538 (0.610–1.295)	0.881/0.576 (0.566–1.373)	0.886/0.503 (0.622–1.262)	1.052/0.790 (0.725–1.525)
Age										
	<63	200/126	295/228	137/112	432/340	1.00	1.285/0.106 (0.948–1.742)	1.384/0.077 (0.965–1.984)	1.316/0.059 (0.990–1.750)	0.845/0.279 (0.623–1.147)
	≥63	185/127	360/275	133/157	493/432	1.00	0.980/0.893 (0.734–1.309)	1.572/0.009 (1.121–2.204)	1.140/0.347 (0.868–1.497)	0.628/0.001 (0.476–0.829)
Smoking status										
	Never	282/149	472/278	206/147	678/425	1.00	1.109/0.417 (0.863–1.426)	1.357/0.042 (1.012–1.821)	1.184/0.159 (0.936–1.499)	0.787/0.057 (0.616–1.007)
	Ever	103/104	183/225	64/122	247/347	1.00	1.088/0.661 (0.746–1.586)	1.816/0.010 (1.154–2.858)	1.272/0.185 (0.891–1.817)	0.582/0.005 (0.399–0.850)
Alcohol consumption										
	Never	358/183	606/335	253/185	859/520	1.00	1.070/0.555 (0.855–1.339)	1.415/0.009 (1.089–1.839)	1.172/0.140 (0.949–1.446)	0.738/0.007 (0.593–0.919)
	Ever	27/70	49/168	17/84	66/252	1.00	1.291/0.378 (0.732–2.276)	1.858/0.083 (0.922–3.741)	1.441/0.184 (0.840–2.471)	0.640/0.138 (0.355–1.154)

Table 6 Stratified Analyses Between Rs8031440 Polymorphism and ESCC Risk by Gender, Age, Smoking Status, and Alcohol Consumption

Variables		Control/Case				Adjusted OR/*P* (95% CI of OR)				
		GG	GA	AA	GA+AA	GG	GA	AA	GA+AA	AA vs (GG+GA)
Gender										
	Male	283/172	474/364	190/209	664/573	1.00	1.200/0.160 (0.931–1.546)	1.805/<0.001 (1.343–2.425)	1.371/0.010 (1.080–1.740)	0.624/<0.001 (0.489–0.796)
	Female	94/78	182/135	86/67	268/202	1.00	0.886/0.530 (0.607–1.292)	0.933/0.757 (0.599–1.451)	0.901/0.565 (0.632–1.285)	0.992/0.966 (0.685–1.436)
Age										
	<63	194/125	299/228	139/113	438/341	1.00	1.238/0.170 (0.913–1.680)	1.353/0.100 (0.944–1.941)	1.274/0.097 (0.957–1.697)	0.845/0.278 (0.624–1.145)
	≥63	183/125	357/271	137/163	494/434	1.00	0.984/0.916 (0.736–1.317)	1.600/0.006 (1.143–2.240)	1.154/0.304 (0.878–1.518)	0.619/0.001 (0.470–0.814)
Smoking status										
	Never	277/146	473/276	210/152	683/428	1.00	1.103/0.445 (0.857–1.421)	1.383/0.030 (1.032–1.854)	1.189/0.152 (0.938–1.507)	0.770/0.036 (0.604–0.983)
	Ever	100/104	183/223	66/124	249/347	1.00	1.055/0.782 (0.723–1.540)	1.742/0.016 (1.109–2.739)	1.232/0.253 (0.861–1.762)	0.595/0.007 (0.409–0.866)
Alcohol consumption										
	Never	351/180	607/333	258/190	865/523	1.00	1.061/0.607 (0.847–1.329)	1.427/0.008 (1.099–1.854)	1.170/0.145 (0.947–1.445)	0.728/0.004 (0.586–0.905)
	Ever	26/70	49/166	18/86	67/252	1.00	1.236/0.467 (0.698–2.186)	1.757/0.112 (0.877–3.519)	1.378/0.247 (0.801–2.371)	0.658/0.156 (0.369–1.174)

$G_{rs8031440}G_{rs8031627}$*SMAD6*$G_{rs3463453}$ (OR=0.809, 95% CI=0.674–0.972, *P*=0.023) and *SMAD3* $A_{rs8028147}G_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$A_{rs3463453}$ (OR=0.569, 95% CI=0.395–0.820, *P*=0.002) were associated with less susceptibility to ESCC, while *SMAD3*$A_{rs8028147}T_{rs3743343}T_{rs3743342}T_{rs8025774}A_{r8031440}A_{rs8031627}$*SMAD6* $A_{rs3463453}$ was associated with higher risk of ESCC (OR = 1.318, 95% CI=1.056–1.645, P= 0.014).

Discussion

Herein, the association between *SMAD3/SMAD6* SNPs and the risk of ESCC among the Chinese population was assessed. Preliminary analysis revealed that the distributions of the 7 SNPs were consistent with that of HapMap data. As such, the results of our study could be generalized and used for the entire Chinese Han population.

Several studies postulate that SMAD3 plays a protective role in ESCC. Our study showed that the association between SMAD3 rs3743342 C>T, rs8025774 C>T, rs8031440 G>A, rs8031627 G>A, and ESCC was consistent in different genetic models. Thus, selected *SMAD3* SNPs may affect the susceptibility of the participants to ESCC in a positive correlated manner. However, mutant heterozygote of the mentioned 4 SNPs seemed to have no significant association with risk of ESCC compared with the wild genotype in the co-dominant model, which may be due to the inheritance mode and mechanism of SNPs on tumor development and progression.

Stratification analyses of the four SNPs further revealed that their effects varied in different subgroups. The SNPs significantly increased the risk of ESCC in males and participants aged more than 63 years. Smokers with *SMAD*3 rs3743342 C>T, rs8025774 C>T, or rs8031627 G>A were more susceptible to ESCC while those with the rs8031440 wild-type homozygotes had a lower risk of ESCC than those with other genotypes. These results were consistent with previous results that showed ESCC is more prevalent in Chinese males than females, and in elder people than younger people, and smoking increases the risk of ESCC by about 3–7-fold.[15,16] Cognizant to this, these results suggested that there exists an interaction between the environmental and genetic risk factors in tumorigenesis of ESCC.

Table 7 Stratified Analyses Between Rs8031627 Polymorphism and ESCC Risk by Gender, Age, Smoking Status, and Alcohol Consumption

Variables		Control/Case				Adjusted OR/P (95% CI of OR)				
		GG	GA	AA	GA+AA	GG	GA	AA	GA+AA	AA vs (GG+GA)
Gender										
	Male	292/173	471/366	185/206	656/572	1.00	1.244/0.090 (0.967–1.602)	1.876/<0.001 (1.396–2.522)	1.421/0.004 (1.120–1.801)	0.614/<0.001 (0.480–0.785)
	Female	95/79	182/136	85/65	267/201	1.00	0.895/0.562 (0.614–1.303)	0.912/0.686 (0.585–1.422)	0.900/0.561 (0.632–1.282)	1.021/0.915 (0.703–1.482)
Age										
	<63	198/125	297/229	137/112	434/341	1.00	1.290/0.101 (0.951–1.749)	1.391/0.074 (0.969–1.996)	1.321/0.056 (0.993–1.759)	0.843/0.274 (0.622–1.145)
	≥63	189/127	356/273	133/159	489/432	1.00	1.003/0.982 (0.752–1.339)	1.628/0.005 (1.162–2.279)	1.173/0.251 (0.894–1.539)	0.616/0.001 (0.467–0.813)
Smoking status										
	Never	282/148	472/278	206/148	678/426	1.00	1.121/0.372 (0.872–1.441)	1.385/0.030 (1.032–1.858)	1.201/0.128 (0.948–1.520)	0.777/0.044 (0.608–0.993)
	Ever	105/104	181/224	64/123	245/347	1.00	1.110/0.585 (0.763–1.617)	1.851/0.008 (1.178–2.909)	1.299/0.148 (0.911–1.853)	0.579/0.005 (0.397–0.844)
Alcohol consumption										
	Never	360/182	604/335	253/186	857/521	1.00	1.089/0.458 (0.870–1.362)	1.446/0.006 (1.113–1.879)	1.194/0.098 (0.968–1.474)	0.730/0.005 (0.586–0.908)
	Ever	27/70	49/167	17/85	66/252	1.00	1.284/0.388 (0.728–2.263)	1.879/0.077 (0.933–3.783)	1.441/0.184 (0.840–2.471)	0.630/0.125 (0.350–1.136)

Table 8 Haplotype Frequencies in the Case and Control Group, and Risk of ESCC

Haplotypes	Case (%)	Control (%)	OR (95% CI)	P
SMAD3 $A_{rs8028147}C_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$ *SMAD6*$A_{rs3463453}$	4.3	4.0	1.075 (0.801–1.443)	0.629
SMAD3 $A_{rs8028147}C_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$G_{rs3463453}$	10.8	12.9	0.809 (0.674–0.972)	0.023
SMAD3 $A_{rs8028147}G_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$A_{rs3463453}$	2.2	3.7	0.569 (0.395–0.820)	0.002
SMAD3 $A_{rs8028147}G_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$G_{rs3463453}$	8.9	9.6	0.912 (0.744–1.117)	0.373
SMAD3 $A_{rs8028147}T_{rs3743343}T_{rs3743342}T_{rs8025774}A_{r8031440}A_{rs8031627}$*SMAD6*$A_{rs3463453}$	8.6	6.7	1.318 (1.056–1.645)	0.014
SMAD3 $A_{rs8028147}T_{rs3743343}T_{rs3743342}T_{rs8025774}A_{rs8031440}A_{rs8031627}$*SMAD6*$G_{rs3463453}$	23.9	21.4	1.148 (0.997–1.321)	0.055
SMAD3 $G_{rs8028147}C_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$A_{rs3463453}$	3.3	4.1	0.779 (0.568–1.067)	0.119
SMAD3 $G_{rs8028147}C_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$G_{rs3463453}$	11.3	10.4	1.100 (0.911–1.328)	0.323
SMAD3 $G_{rs8028147}T_{rs3743343}C_{rs3743342}C_{rs8025774}G_{rs8031440}G_{rs8031627}$*SMAD6*$G_{rs3463453}$	5.9	6.7	0.872 (0.684–1.111)	0.267
SMAD3 $G_{rs8028147}T_{rs3743343}T_{rs3743342}T_{rs8025774}A_{rs8031440}A_{rs8031627}$*SMAD6*$A_{rs3463453}$	4.9	4.4	1.117 (0.846–1.476)	0.434
SMAD3 $G_{rs8028147}T_{rs3743343}T_{rs3743342}T_{rs8025774}A_{rs8031440}A_{rs8031627}$*SMAD6*$G_{rs3463453}$	13.1	12.9	1.018 (0.855–1.213)	0.838

Non-alcohol drinking participants with *SMAD3* rs3743342 C>T, rs8025774 C>T, rs8031440, or rs8031627 G>A were more susceptible to ESCC. However, there was no such correlation in the alcohol-drinking subgroup. The result seemed contradictory to the evidence that alcohol drinking is a significant contributory factor to the development of ESCC.[17] Thus, the mechanism underlying this discrepancy should be further investigated.

There was a significant correlation between rs3743342, rs8025774, rs8031440, and rs8031627 that further confirmed their similarities. The four SNPs were all located in the three-prime untranslated region (3′-utr) of *SMAD3*. Many recent studies have reported that the 3′-utr of *SMAD3* has an important impact on the development of various malignant tumors. For example, inhibition of microRNAs that target at the 3′-utr of *SMAD3* leads to the

upregulation of *SMAD3*, thereby constraining the epithelial–mesenchymal transition and invasion of non-small cell lung cancer.[18] In the same line, silencing the micro-RNAs that target the 3′-utr of *SMAD3* decreases the expression of *SMAD3*, thereby inhibiting the proliferation of glioblastoma cells.[19] Based on these reports, it appears that these SNPs influence the risk of ESCC through post-transcriptional regulation. Raine et al reported that rs8031440 and rs3743342 were also correlated with primary osteoarthritis, aneurysms, and osteoarthritis syndrome.[20] It can be reasonably assumed that these SNPs affect the susceptibility to ESCC as well as that to other diseases.

The major limitation of our study was the lack of technical support to establish a single nucleotide mutation cell or animal model. As such, the biological function of these SNPs requires further research. In addition, our study was conducted in a single center, although the sample size was impressive.

Conclusion

SMAD3 rs3743342 C>T, rs8025774 C>T and rs8031627 G>A increase the susceptibility of individuals to ESCC, particularly in males, people aged over 63 years, smokers, and non-alcohol drinking people. The distribution of the 7 SNPs was consistent with that of HapMap data based on their primary data. As such, the results of this study can be generalized and used as a useful resource for ESCC screening of the entire Chinse Han population.

Author Contributions

All authors contributed to data analysis, drafting or revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

All authors have no conflicts of interest to declare.

References

1. Budi EH, Duan D, Derynck R. Transforming growth factor-β receptors and SMADs: regulatory complexity and functional versatility. *Trends Cell Biol*. 2017;27:658–672. doi:10.1016/j.tcb.2017.04.005
2. Macias MJ, Martin-Malpartida P, Massagué J. Structural determinants of SMAD function in TGF-β signaling. *Trends Biochem Sci*. 2015;40:296–308. doi:10.1016/j.tibs.2015.03.012
3. Miyazawa K, Miyazono K. Regulation of TGF-β family signaling by inhibitory SMADs. *CSH Perspect Biol*. 2017;9:a22095.
4. Chen W, Ten Dijke P. Immunoregulation by members of the TGFβ superfamily. *Nat Rev Immunol*. 2016;16:723–740. doi:10.1038/nri.2016.112
5. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov*. 2019;18:197–218. doi:10.1038/s41573-018-0007-y
6. Zhu H, Gu X, Xia L, et al. A novel TGFβ trap blocks chemotherapeutics-induced TGFβ1 signaling and enhances their anticancer activity in gynecologic cancers. *Clin Cancer Res*. 2018;24:2780–2793. doi:10.1158/1078-0432.CCR-17-3112
7. Fang X, Gao Y, Li Q. SMAD3 activation: a converging point of dysregulated TGF-beta superfamily signaling and genetic aberrations in granulosa cell tumor development? *Biol Reprod*. 2016;95:105. doi:10.1095/biolreprod.116.143412
8. Matsuzaki K, Kitano C, Murata M, et al. SMAD2 and SMAD3 phosphorylated at both linker and COOH-terminal regions transmit malignant TGF-β signal in later stages of human colorectal cancer. *Cancer Res*. 2009;69:5321–5330. doi:10.1158/0008-5472.CAN-08-4203
9. Bae E, Sato M, Kim RJ, et al. Definition of SMAD3 phosphorylation events that affect malignant and metastatic behaviors in breast cancer cells. *Cancer Res*. 2014;74:6139–6149. doi:10.1158/0008-5472.CAN-14-0803
10. Zhao Y, Ma J, Fan Y, et al. TGF-beta transactivates EGFR and facilitates breast cancer migration and invasion through canonical SMAD3 and ERK/Sp1 signaling pathways. *Mol Oncol*. 2018;12:305–321. doi:10.1002/1878-0261.12162
11. Han S, Kim H, Seong DH, et al. Loss of the SMAD3 expression increases susceptibility to tumorigenicity in human gastric cancer. *Oncogene*. 2004;23:1333–1341. doi:10.1038/sj.onc.1207259
12. Cho SY, Ha SY, Huang S, et al. The prognostic significance of SMAD3, SMAD4, SMAD3 phosphoisoform expression in esophageal squamous cell carcinoma. *Med Oncol*. 2014;31. doi:10.1007/s12032-014-0236-9
13. Jeon HS, Dracheva T, Yang SH, et al. SMAD6 contributes to patient survival in non-small cell lung cancer and its knockdown reestablishes TGF- homeostasis in lung cancer cells. *Cancer Res*. 2008;68:9686–9692. doi:10.1158/0008-5472.CAN-08-1083
14. Yin J, Wang L, Tang W, et al. RANK rs1805034 T>C polymorphism is associated with susceptibility of esophageal cancer in a Chinese population. *PLoS One*. 2014;9:e101705. doi:10.1371/journal.pone.0101705
15. Zheng R, Zeng H, Zhang S, Chen T, Chen W. National estimates of cancer prevalence in China, 2011. *Cancer Lett*. 2016;370:33–38. doi:10.1016/j.canlet.2015.10.003
16. Kamangar F, Chow W, Abnet C, Dawsey S. Environmental causes of esophageal cancer. *Gastroenterol Clin North Am*. 2009;38:27–57. doi:10.1016/j.gtc.2009.01.004
17. Salaspuro MP. Alcohol consumption and cancer of the gastrointestinal tract. *Best Pract Res Clin Gastroenterol*. 2003;17:679–694. doi:10.1016/S1521-6918(03)00035-0
18. Hu H, Xu Z, Li C, et al. MiR-145 and miR-203 represses TGF-β-induced epithelial-mesenchymal transition and invasion by inhibiting SMAD3 in non-small cell lung cancer cells. *Lung Cancer*. 2016;97:87–94. doi:10.1016/j.lungcan.2016.04.017
19. Wu ZB, Cai L, Lin SJ, Lu JL, Yao Y, Zhou LF. The miR-92b functions as a potential oncogene by targeting on SMAD3 in glioblastomas. *Brain Res*. 2013;1529:16–25. doi:10.1016/j.brainres.2013.07.031
20. Raine EVA, Reynard LN, van de Laar IMBH, Bertoli-Avella AM, Loughlin J. Identification and analysis of a SMAD3 cis-acting eQTL operating in primary osteoarthritis and in the aneurysms and osteoarthritis syndrome. *Osteoarthritis Cartil*. 2014;22:698–705. doi:10.1016/j.joca.2014.02.931