Open Access Full Text Article

ORIGINAL RESEARCH

The Single Nucleotide Polymorphisms of AP1S1 are Associated with Risk of Esophageal Squamous Cell Carcinoma in Chinese Population

Feng Su[1,*], Yong Fang [1,*], Jinjie Yu [1], Tian Jiang[1], Siyun Lin[1], Shaoyuan Zhang [1], Lu Lv[2], Tao Long[2], Huiwen Pan[2], Junqing Qi[2], Qiang Zhou[3], Weifeng Tang[4], Guowen Ding[2], Liming Wang[5], Lijie Tan[1], Jun Yin[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; [3]Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan, People's Republic of China; [4]Department of Cardiothoracic Surgery, Nanjing Drum Tower Hospital, Jiangsu, People's Republic of China; [5]Department of Respiratory, Shanghai Xuhui Central Hospital, Shanghai, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jun Yin; Lijie Tan, Zhongshan Hospital of Fudan University, 180 Fenglin road, Xuhui District, Shanghai, 200032, People's Republic of China, Email jun_yin@fudan.edu.cn; tan.lijie@zs-hospital.sh.cn

Background: The σ1A subunit of the adaptor protein 1 (AP1S1) participates in various intracellular transport pathways, especially the maintenance of copper homeostasis, which is pivotal in carcinogenesis. It is therefore rational to presume that AP1S1 might also be involved in carcinogenesis. In this hospital-based case-control study, we investigated the genetic susceptibility to ESCC in relation to SNPs of *AP1S1* among Chinese population.

Methods: A database containing a total of 1303 controls and 1043 ESCC patients were retrospectively studied. The *AP1S1* SNPs were analyzed based on ligation detection reaction (LDR) method. Then, the relationship between ESCC and SNPs of AP1S1 was determined with a significant crude P<0.05. Then the logistic regression analysis was used for the calculation for adjusted P in the demographic stratification comparison if a significant difference was observed in the previous step.

Results: *AP1S1* rs77387752 C>T genotype TT was an independent risk factor for ESCC, while rs4729666 C>T genotype TC and rs35208462 C>T genotype TC were associated with a lower risk for ESCC, especially in co-dominant model and allelic test for younger, male subjects who are not alcohol-drinkers nor cigarette smokers.

Conclusion: *AP1S1* rs77387752, rs4729666 and rs35208462 polymorphisms are associated with susceptibility to ESCC in Chinese individuals. *AP1S1* SNPs may exert an important role in esophageal carcinogenesis and could serve as potential diagnostic biomarkers.

Keywords: AP1S1, single nucleotide polymorphism, SNP, esophageal squamous cell carcinoma, ESCC

Introduction

Esophageal cancer is the 6th most fatal malignancy with 500,000 new cases diagnosed annually, and remains a major threat to public health.[1,2] Esophageal squamous cell carcinoma (ESCC) is the predominant histological type, with a global incidence of 5.2/100,000.[3] Despite the tremendous advances in multidisciplinary therapy, the prognosis is still dismal with 5-year overall survival rate ranging from 15~25%, and the incidence of esophageal cancer is still rising.[4,5] It is generally acknowledged that the occurrence of esophageal cancer is multifactorial, senility, gender and environmental factors such as alcohol consumption,[6,7] tobacco use,[8] poor oral hygiene, nutritional deficiencies[9–11] could jointly contribute to the carcinogenesis.[12] But only part of individuals exposed to these risk factors finally develop esophageal cancer, suggesting that genetic susceptibility including single nucleotide polymorphisms (SNPs) is important for esophageal carcinogenesis.

The clathrin is linked to receptors in coated vesicles by clathrin coat assembly complex, and the σ1A small subunit of the adaptor protein (AP) 1 complex (AP1S1) is part of the coat assembly complex. As a part of endocytosis and Golgi processing, AP coordinates various transportations of intracellular endomembrane pathway.[13] The *AP1S1* defects were correlated with abnormal subcellular localization of two Golgi-localized copper transport ATPases ATP7A and ATP7B,[14] causing copper homeostasis disorder. Harboring the two copper transports above, the Golgi apparatus regulates the copper supply of several carcinogenic metalloenzymes, and thus governs the overall utilization of copper in carcinoma. Accumulating evidence has suggested that copper is involved in various cancer-associated processes,[15] including mitogen-activated kinases signaling pathway,[16,17] balance of oxidation with antioxidant balance,[18] cellular respiration[19] and the formation of covalent bonds of extracellular matrix proteins.[20] In addition, the generation of reactive oxygen species (ROS) including hydroxyl radical and superoxide radical, which can break the DNA strands or modify the bases causing carcinogenesis,[21] is closely related to copper. Oxidative stress caused by copper led to mitochondrial-mediated apoptosis and autophagy,[22–24] which has a context-dependent role in cancer. Furthermore, inhibitory effects on tumor growth and malignancy of copper chelators have also been documented.[25,26] Indeed, patients with Hodgkin's disease have a higher serum copper level, which is strongly correlated with the histopathological types, clinical stage and prognosis.[27] Moreover, copper concentration in cancerous tissues of patients with large bowel cancer and esophageal cancer was lower.[28] Similarly, the expressions and SNPs of copper transport ATPases ATP7A and ATP7B, are associated with the response to chemotherapy and survival in patients with non-small cell lung cancer.[29,30] Given that AP1S1 can direct the intracellular transportation of copper pumps, affect the subcellular localization of copper transport ATPases ATP7A/7B and regulate the intracellular copper level, it is logical to speculate that AP1S1 plays a critical role in carcinogenesis. Consistently, in acute myeloid leukemia and prostate carcinoma gene expression and methylation profiles, an association rule mining based study identified *AP1S1* as an integrated marker through mutual information based statistically significant feature extraction.[31]

However, the correlation between AP1S1 and ESCC remains obscure, and the effect of *AP1S1* SNPs on ESCC is unclear. Thus, we conducted a hospital-based, hypothesis-driven, case-control study to assess the molecular mechanisms underlying the functional *AP1S1* SNPs in ESCC. To further explore the ingredient between the SNPs and ESCC in certain population, we did subgroup analysis selected identified typical risk factors including gender, age, drinking and smoking status for subgroup analysis.

Methods
Ethics Statement
The protocol conforms to the *World Medical Association Declaration of Helsinki* on the research ethical behavior of human/animal subjects and was approved by the Review Board of Jiangsu University (Zhenjiang, China). All participants were provided with written informed consent.

Study Subjects
A total of 2186 participants were retrospectively enrolled from the Affiliated People's Hospital and Affiliated Hospital of Jiangsu University (Zhenjiang, China) between October 2008 and January 2017. A total of 1043 cases of ESCC were diagnosed and confirmed histologically. Patients with any type of cancer history/metastasized cancer or had the history of neoadjuvant therapy were excluded for the study. Totally, a total of 1143 non-cancer individuals were enrolled from two hospitals around the same time period with frequency matched to cases in regard to age (± 5 years) and gender, most of which were admitted to the hospital out of trauma.

The demographic characteristics such as age, sex and other risk factors for ESCC including smoking and drinking status were collected in 1043 patients and 1315 negative controls by questionnaire. Feedback was obtained from 1143 control subjects and all case subjects. 2 mL blood samples were collected intravenously in advance from each participant for subsequent analysis. Smoker subgroup was defined as 1 cigarette per day for more than one year with or without a cessation history, while alcohol drinker subgroup was defined as more than 3 alcoholic drinks per week for more than six months.

Genomic DNA Extraction and SNPs Analysis

Genomic DNA was isolated from peripheral blood with QIAamp DNA Blood Mini Kit (Qiagen, Berlin, Germany) and amplified by PCR.[32] The samples were further genotyped using the ligation detection reaction (LDR) method (technical support from Genesky Biotechnology Inc. Shanghai, China). *AP1S1* rs6969685, rs4727480, rs77387752, rs4729666 and rs35208462 SNPs were selected and analyzed.[33,34] Quality control was conducted by repeating the analyses in 10% of randomly selected samples. In pilot study, the linkage disequilibrium analysis was done using 1000Genomes database to select the SNP loci with correlation, and then the tag SNPs were further analyzed.

Statistical Analysis

The statistical software package SPSS 20.0 (SPCC Inc., Chicago, IL) was used for the analysis of the demographic distribution, Hardy–Weinberg equilibrium (HWE) and genotype for cases and controls with χ^2 or Fisher's exact good-ness-of-fit test according to different applicable conditions. The crude odds ratio (OR) and the corresponding 95% confidence intervals (CIs) were calculated with the genotype between two arms and were used to compared the risk with parametric test in most cases and nonparametric if the sample size in the group was small. Then the adjusted OR and corresponding CIs were calculated with logistics regression analysis for hierarchical analysis based on demographic information for the SNPs considered relevant to ESCC in the primary results. The demographics including age, gender, alcohol, and tobacco consuming status were covariances, while the genotype was dummy variable and the group was dependent variable in the logistics for subgroup analysis. Statistical significance is considered when a two-tailed P value of < 0.05 is observed, while the two-tailed *P* value between 0.05 and 0.10 was judged as borderline statistically significant. Then we tested the linkage disequilibrium analyses using SHEsis online platform[35] and then accessed OR of cases and controls in haplotypes by SHEsis as well.[36]

Results

Characteristics of Study Subjects

The demographic characteristics of the 1043 ESCC patients and 1143 controls were summarized in Table 1. Using χ^2 tests, we found that was no significant difference in age and sex between the case and control groups (P=0.764, P=0.903), while the ESCC group has a significant higher rate of smoking and alcohol drinking history (P<0.001).

The primary information of 5 genotyped SNPs of *AP1S1* was provided in Table 2. The success rates of SNP genotyping exceeded 95%. Minor allele frequencies (MAFs) in the controls were comparable to global and East Asian MAFs in the 1000 Genomes database. Deviation tests for the Hardy-Weinberg (HWE) revealed that the control group was in the Hardy-Weinberg proportions for all the 5 SNPs with P value higher than 0.05.

Association Analyses of *AP1S1* SNPs with ESCC

The association between 5 *AP1S1* SNPs and ESCC was shown in Table 3. In the co-dominant test, rs77387752 C>T genotype TT was a risk factor for ESCC (P=0.028, OR=10.218), while the genotype TC for *AP1S1* rs4729666 C>T (P=0.008, OR=0.751) and rs35208462 C>T (P=0.021, OR=0.780) were associated with a lower risk for ESCC. In the dominant model test, rs4729666 C>T (P=0.011, OR=0.766) and rs35208462 C>T (P=0.012, OR=0.767) were associated reduced risk of ESCC. In the recessive model test, rs77387752 C>T was associated with a significantly higher risk for ESCC (P=0.029, OR=10.072). In the allelic test, marginal statistical significance was observed in genotype frequencies of *AP1S1* rs77387752 C>T SNP (P=0.067, OR=1.255), while rs4729666 C>T (P=0.024, OR=0.806) and rs35208462 C>T (P=0.009, OR=0.777) have significant difference.

Stratification Risk of ESCC Associated with SNPs

Next, we conducted stratification analyses of rs77387752, rs4729666 and rs35208462 according to demographic information to further access the role of AP1S1 on the risk of ESCC (Tables 4–6). In the non-smokers, rs77387752 C>T, a marginal significance was observed in the recessive model (P=0.054, adjusted OR=8.060) and in the co-dominant model when compared TT with CC (P=0.054, adjusted OR=8.045), while cigarette smoking is a risky factor in the

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

Variable	Control (n=1143)	Case (n=1043)	P
	n (%)	n (%)	
Age (Years)			
Mean (SD)	62.64 (9.902)	63.07 (7.271)	0.252
<65	644 (56.3)	581(55.7)	0.764
≥65	499 (43.7)	462(44.3)	
Gender			
Male	828 (72.4)	758(72.7)	0.903
Female	315 (27.6)	285(27.3)	
Smoking status			
Never	810 (70.9)	589(56.5)	**<0.001**
Ever	333 (29.1)	454(43.5)	
Alcohol consumption			
Never	1061 (92.8)	714(68.5)	**<0.001**
Ever	82 (7.2)	329(31.5)	

Note: Bold font indicates significance at the p<0.05 level.

Table 2 Primary Information for AP1S1 rs6969685 T>C, rs4727480 A>G, rs77387752 C>T, rs4729666 C>T, rs35208462 C>T

Genotyped SNP	rs6969685	rs4727480	rs77387752	rs4729666	rs35208462
Allele	T>C	A>G	C>T	C>T	C>T
Gene	AP1S1				
Consequence	Unreported	Intron variant	Synonymous variant	Intron variant	Intron variant
Regulome DB score*	5	4	4	2a	2b
Chromosome	7				
Chromosome position	101,152,465	101,156,292	101,156,689	101,159,880	101,160,408
MAF in 1000 Genomes					
Global	C=0.4946	G=0.2927	T=0.0198	T=0.1086	T=0.1480
East Asian	C=0.4732	G=0.1448	T=0.0744	T=0.1171	T=0.1121
MAF in control	C=0.4697	G=0.156	T=0.0580	T=0.128	T=0.128
P value for HWE test in the controls	0.996	0.523	0.125	0.646	0.646
Genotyping method	LDR	LDR	LDR	LDR	LDR
% Genotyping value	98.90%	98.90%	98.90%	98.20%	98.20%

Note: *http://www.regulomedb.org/.
Abbreviation: LDR, ligation detection reaction.

Table 3 Logistic Regression Analyses of Associations Between rs6969685 T>C, rs4727480 A>G, rs77387752 C>T, rs4729666 C>T, rs35208462 C>T and Risk of ESCC

Locus	Genotype	Control	Case	Co-Dominant Model		Dominant Model	Recessive Model	Additive Model
				OR (95% CI)	P	OR (95% CI)/P	OR (95% CI)/P	OR (95% CI)/P
rs6969685	TT	320	283	1	0.957	1.026	1.022	1.018
	TC	567	512	1.021(0.836–1.247)	0.838	(0.850–1.238)	(0.835–1.252)	(0.903–1.147)
	CC	251	230	1.036(0.815–1.317)	0.772	P=0.792	P=0.831	P=0.769
rs4727480	AA	807	734	1	0.771	0.967	1.159	0.987
	AG	306	265	0.952(0.785–1.154)	0.617	(0.802–1.165)	(0.665–2.202)	(0.837–1.163)
	GG	25	26	1.143(0.654–1.998)	0.638	P=0.721	P=0.603	P=0.872
rs77387752	CC	1007	887	1	0.061	1.196	10.072	1.255
	CT	130	129	1.127(0.869–1.461)	0.369	(0.926–1.544)	(1.274–79.635)	(0.984–1.600)
	TT	1	9	10.218(1.292–80.807)	**0.028**	P=0.170	**P=0.029**	P=0.067
rs4729666	CC	863	811	1	**0.030**	0.766	1.062	0.806
	CT	258	182	0.751(0.607–0.928)	**0.008**	(0.623–0.942)	(0.534–2.114)	(0.668–0.972)
	TT	17	16	1.002(0.503–1.996)	0.997	**P=0.011**	P=0.863	**P=0.024**
rs35208462	CC	863	810	1	**0.032**	0.767	0.594	0.777
	CT	258	189	0.780(0.632–0.964)	**0.021**	(0.624–0.943)	(0.264–1.339)	(0.644–0.939)
	TT	17	9	0.564(0.250–1.273)	0.168	**P=0.012**	P=0.209	**P=0.009**

Note: Bold font indicates significance at the p<0.05 level.

dominant model (P=0.087, adjusted OR=1.519). In the non-drinking subjects, rs77387752 C>T was statically significant in the recessive model (P=0.033, adjusted OR=9.816) and in the co-dominant model when compared TT with CC (P=0.032, adjusted OR=9.924). In the male (P=0.001, adjusted OR=0.624), participants younger than 65 (P=0.016, adjusted OR=0.691), smoking participants (P=0.002, adjusted OR=0.418) and no alcohol drinking (P=0.054, adjusted OR=0.788) participants rs4729666 C>T TC was associated lower risk of ESCC compared with CC, and in the dominant model test rs4729666 C>T had a lower likelihood of having ESCC (P=0.002/0.027/0.004/0.073, adjusted OR=0.657/0.721/0.565/0.808), while we got a similar result in the participants younger than 65, non-smokers and non-alcohol drinking participants in the rs35208462 C>T in the dominant model test (P=0.086/0.019/0.044, adjusted OR=0.774/0.732/0.785). Compared with CC, the genotype TC in rs35208462 C>T was associated lower risk for ESCC in non-smokers (P=0.029, adjusted OR=0.774) and non-drinkers (P=0.066, adjusted OR=0.798). However, the mutant homozygote in specific stratified groups of rs77387752 C>T and rs4729666 C>T were zero, the P and OR in these groups were incalculable.

Haplotype Polymorphisms and Susceptibility to ESCC

Next, we used SHEsis online platform and summarized the haplotype analysis of 5 SNPs in Tables 7–9. *AP1S1* $C_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}C_{rs35208462}$ (P=0.0107, OR=1.177 95% CI=1.038–1.334) and *AP1S1* $T_{rs6969685}G_{rs4727480}C_{rs77387752}C_{rs4729666}C_{rs35208462}$ (P<0.001, OR=1.922, 95% CI=1.395–2.674) are associated with high susceptibility to ESCC, while *AP1S1* $C_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}T_{rs35208462}$ (P=0.0028, OR=0.747, 95% CI=0.616–0.905) and *AP1S1* $T_{rs6969685}G_{rs4727480}C_{rs77387752}T_{rs4729666}C_{rs35208462}$ (P=0.0087, OR=0.776, 95% CI=0.641–0.938) were associated with lower risk of ESCC.

Discussion

In this hospital-based case-control study, we identified *AP1S1* rs77387752 C>T genotype TT was an independent risk factor for ESCC, while rs4729666 C>T genotype TC and rs35208462 C>T genotype TC were associated with a lower risk for ESCC, especially in younger, male cohort with alcohol but with cigarette consuming. To our knowledge, this is the first report demonstrating the association between *AP1S1* SNPs and ESCC susceptibility.

Table 4 Stratified Analyses Between rs77387752 C>T Polymorphism and ESCC Risk by Sex, Age, Smoking Status, and Alcohol Consumption

Variables	Control/Case					Adjusted OR/ (95% CI of OR) /P				
Genotype	CC	TC	TT	CC+TC	TC+TT	CC	TC	TT	CC VS (TC+TT)	(CC+TC) vs TT
Gender										
Male	727/642	97/97	0/6	824/739	97/103	-/-/1.000	1.154/ (0.835–1.595) 0.385	-/ -/ -	1.234/(0.896–1.697)/0.197	-/ -/ -
Female	280/245	33/32	1/3	313/277	34/35	-/-/1.000	1.033/ (0.612–1.742)/0.903	3.501/(0.358–34.232)/0.281	1.103/(0.663–1.835)/0.704	3.489/(0.357–34.092)/0.283
Age										
<65	568/503	71/69	1/3	639/572	72/72	-/-/1.000	1.218/ (0.837–1.771)/0.302	2.141/(0.189–24.271)/0.539	1.234/(0.852–1.787)/0.266	2.107/(0.186–23.888)/0.547
≥65	439/384	59/60	0/6	498/444	59/66	-/-/1.000	1.026/ (0.684–1.541)/0.900	-/ -/ -	1.165/(0.784–1.731)/0.451	-/ -/ -
Smoking status										
Never	711/501	94/67	1/6	805/568	95/73	-/-/1.000	1.015/ (0.724–1.424)/0.931	8.060/(0.963–67.482)/0.054	1.093/(0.786–1.521)/0.598	8.045/(0.961–67.334)/0.054
Ever	296/386	36/62	0/3	332/448	36/65	-/-/1.000	1.454/ (0.897–2.357)/0.129	-/ -/ -	1.519/(0.941–2.454)/0.087	-/ -/ -
Alcohol consumption										
Never	932/609	123/87	1/7	1055/696	124/94	-/-/1.000	1.096/ (0.816–1.472)/0.542	9.924/(1.211–81.294)/**0.032**	1.171/(0.877–1.563)/0.283	9.816/(1.199–80.384)/ **0.033**
Ever	75/278	7/42	0/2	82/320	7/44	-/-/1.000	1.691/ (0.721–3.968)/0.227	-/ -/ -	1.759/(0.752–4.113)/0.193	-/ -/ -

Note: Bold font indicates significance at the p<0.05 level.

Table 5 Stratified Analyses Between Rs4729666 C>T Polymorphism and ESCC Risk by Sex, Age, Smoking Status, and Alcohol Consumption

Variables	Control/Case					Adjusted OR/(95% CI of OR)/P				
Genotype	CC	TC	TT	CC+TC	TC+TT	CC	TC	TT	CC VS (TC+TT)	(CC+TC) vs TT
Gender										
Male	623/603	192/118	9/10	815/721	201/128	-/-/1.000	0.624/(0.473–0.823)/**0.001**	1.397/(0.540–3.610)/0.490	0.657/(0.502–0.859)/**0.002**	1.534/(0.594–3.597)/0.376
Female	240/208	66/64	8/6	306/272	74/70	-/-/1.000	1.160/(0.781–1.723)/0.463	1.016/(0.343–3.006)/0.977	1.145/(0.782–1.677)/0.485	0.980/(0.333–2.887)/0.971
Age										
<65	472/451	155/100	13/12	627/551	72/72	-/-/1.000	0.691/(0.510–0.934)/**0.016**	1.089/(0.476–2.489)/0.840	0.721/(0.539–0.964)/**0.027**	1.185/(0.520–2.702)/0.686
≥65	391/360	103/82	4/4	494/442	59/66	-/-/1.000	0.861/(0.614–1.210)/0.389	1.372/(0.333–5.664)/0.662	0.879/(0.630–1.227)/0.449	1.414/(0.343–5.827)/0.632
Smoking status										
Never	621/445	171/111	14/11	792/556	185/121	-/-/1.000	0.894/(0.682–1.173)/0.419	1.061/(0.473–2.379)/0.886	0.907/(0.698–1.179)/0.465	1.087/(0.485–2.432)/0.840
Ever	242/366	87/71	3/5	329/437	90/76	-/-/1.000	0.535/(0.359–0.795)/**0.002**	1.559/(0.337–7.205)/0.570	0.565/(0.383–0.832)/**0.004**	1.802/(0.391–8.309)/0.450
Alcohol consumption										
Never	801/550	238/130	17/13	1039/680	255/143	-/-/1.000	0.788/(0.619–1.004)/0.054	1.076/(0.515–2.247)/0.845	0.808/(0.639–1.020)/0.073	1.134/(0.544–2.362)/0.738
Ever	62/261	20/52	0/3	82/313	20/55	-/-/1.000	0.627/(0.347–1.132)/0.122	-/ -/ -	0.666/(0.370–1.199)/0.175	-/ -/ -

Note: Bold font indicates significance at the p<0.05 level.

Table 6 Stratified Analyses Between Rs35208462 C>T Polymorphism and ESCC Risk by Sex, Age, Smoking Status, and Alcohol Consumption

Variables	Control/Case					Adjusted OR/(95% CI of OR)/P				
Genotype	CC	TC	TT	CC+TC	TC+TT	CC	TC	TT	CC VS (TC+TT)	(CC+TC) vs TT
Gender										
Male	626/589	186/137	12/7	812/726	198/144	-/-/1.000	0.822/(0.629–1.075)/ 0.152	0.563/(0.202–1.572)/0.273	0.805/(0.619–1.047)/0.106	0.586/(0.210–1.634)/0.307
Female	237/221	72/52	5/2	309/273	77/54	-/-/1.000	0.776/(0.516–1.167)/0.223	0.478/(0.091–2.512)/0.383	0.757/(0.508–1.130)/ 0.173	0.505/(0.096–2.648)/0.419
Age										
<65	479/449	151/107	10/5	630/556	161/112	-/-/1.000	0.788/(0.584–1.062)/0.117	0.572/(0.182–1.800)/0.340	0.774/(0.578–1.037)/ 0.086	0.603/(0.192–1.893)/0.386
≥65	384/361	107/82	7/4	491/443	114/86	-/-/1.000	0.838/(0.597–1.175)/0.304	0.507/(0.134–1.918)/0.317	0.816/ (0.586–1.136)/ 0.228	0.525/(0.139–1.982)/0.342
Smoking status										
Never	594/451	201/111	11/4	795/562	212/115	-/-/1.000	0.774/(0.571–0.970)/**0.029**	0.506/ (0.159–1.609)/ 0.248	0.732/(0.564–0.950)/ **0.019**	0.541/(0.170–1.717)/0.297
Ever	269/359	57/78	6/5	326/437	63/83	-/-/1.000	0.951/(0.622–1.452)/ 0.815	0.564/(0.147–2.167)/0.404	0.913/(0.606–1.375)/0.662	0.569/(0.148–2.181)/0.411
Alcohol consumption										
Never	800/554	241/132	15/6	1041/686	254/138	-/-/1.000	0.798/(0.628–1.015)/0.066	0.581/(0.223–1.513)/0.266	0.785/(0.621–0.994)/**0.044**	0.609/(0.234–1.585)/0.310
Ever	63/256	17/57	2/3	80/313	19/60	-/-/1.000	0.851/(0.459–1.578)/0.610	0.341/(0.055–2.101)/0.246	0.795/(0.439–1.438)/0.447	0.351/(0.057–2.157)/0.258

Note: Bold font indicates significance at the p<0.05 level.

Table 7 Haplotype Frequencies in the Case and Control Group, and Risk of ESCC

Haplotype	Chi2	Case (%)	Control (%)	OR (95% CI)	P
AP1S1 $C_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}C_{rs35208462}$	6.510	37.8	34.3	1.177 (1.038–1.334)	0.0107
AP1S1 $C_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}T_{rs35208462}$	8.877	9.7	12.7	0.747 (0.616–0.905)	**0.0028**
AP1S1 $T_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}C_{rs35208462}$	1.596	29.6	31.6	0.919 (0.807–1.048)	0.2064
AP1S1 $T_{rs6969685}A_{rs4727480}C_{rs77387752}C_{rs4729666}T_{rs35208462}$	–	0.2	0	–	–
AP1S1 $T_{rs6969685}A_{rs4727480}T_{rs77387752}C_{rs4729666}C_{rs35208462}$	2.171	6.9	5.8	1.204 (0.940–1.541)	0.1406
AP1S1 $T_{rs6969685}G_{rs4727480}C_{rs77387752}C_{rs4729666}C_{rs35208462}$	16.485	5.1	2.7	1.922 (1.395–2.674)	**<0.001**
AP1S1 $T_{rs6969685}G_{rs4727480}C_{rs77387752}C_{rs4729666}T_{rs35208462}$	–	0	0.1	–	–
AP1S1 $T_{rs6969685}G_{rs4727480}C_{rs77387752}T_{rs4729666}C_{rs35208462}$	6.871	10.01	12.8	0.776 (0.641–0.938)	**0.0087**
AP1S1 $T_{rs6969685}G_{rs4727480}C_{rs77387752}T_{rs4729666}T_{rs35208462}$	–	0.1	0.1	–	–
AP1S1 $C_{rs6969685}A_{rs4727480}C_{rs77387752}T_{rs4729666}C_{rs35208462}$	–	0.2	0	–	–
AP1S1 $T_{rs6969685}A_{rs4727480}C_{rs77387752}T_{rs4729666}C_{rs35208462}$	–	0.1	0	–	–
AP1S1 $T_{rs6969685}A_{rs4727480}T_{rs77387752}C_{rs4729666}T_{rs35208462}$	–	0.1	0	–	–
AP1S1 $T_{rs6969685}A_{rs4727480}T_{rs77387752}T_{rs4729666}C_{rs35208462}$	–	0	0	–	–
AP1S1 $T_{rs6969685}G_{rs4727480}T_{rs77387752}C_{rs4729666}C_{rs35208462}$	–	0.1	0	–	–

Note: Bold font indicates significance at the p<0.05 level.

Table 8 Linkage Disequilibrium Analysis Using Linkage Disequilibrium Parameter D

D'	rs4727480	rs77387752	rs4729666	rs35208462
rs6969685	1.000	0.999	0.978	0.959
rs4727480	–	0.821	0.981	0.944
rs77387752	–	–	0.899	0.891
rs4729666	–	–	–	0.920

Table 9 Linkage Disequilibrium Analysis Using Linkage Disequilibrium Parameter r^2

r^2	rs4727480	rs77387752	rs4729666	rs35208462
rs6969685	0.164	0.062	0.115	0.135
rs4727480	–	**0.009**	0.698	**0.022**
rs77387752	–	–	**0.007**	**0.007**
rs4729666	–	–	–	**0.015**

Note: Bold font indicates significance at the p<0.05 level.

AP-1 complex is an oncogene factor in the Hippo pathway mediate the phosphorylation and inactivation of YAP and TAZ and it is associated with development and tissue repair.[37] Pervasive activation of the pathway is associated with a multitude of human tumors and the acquisition of malignant traits.[38] As a subunit of AP1, AP1S1, a peptide chain of 121 amino acids, has been reported associated with MEDNIK syndrome.[14,39–41] The *AP1S1* defects were correlated with

abnormal subcellular localization of copper transport ATPases ATP7A and ATP7B,[14] leading to copper homeostasis disorder which is critical in carcinogenesis. Based on the above evidence, we sought to investigate the role of AP1S1 in ESCC and seek the association between *AP1S1* SNPs and ESCC.

In this study, we showed significant association between *AP1S1* rs35208462 C>T, rs4729666 C>T, rs77387752 C>T and ESCC in certain genotypes and test models. However, we observed a protective effect in the heterozygous mutation of rs35208462 C>T and rs4729666 C>T but not in the homozygous mutation. This finding is similar to previous report that different phenotypes existed in homozygous and heterozygous mutations of TNFRSF11A in mice.[42] With a P value of 0.168 in homozygous mutation rs35208462 C>T, the insufficient number of participants may also compromise the power of statistical analysis. The relationship between ESCC and mutant heterozygote of rs77387752 is not statically significant as well, which may be due to the mode of inheritance and mechanism of SNPs in tumorigenesis and development.

The MAF of rs4727480 and rs77387752 in the control group are quite different from previously reported in the database. This may be due to fact that we consulted the database containing subjects in the worldwide, while our study and control cohorts were from the east part of China. This heterogeneity of allele frequency between our control group and database might reflect the regional characteristic of natural selection and infection pressure.[43]

Stratification analyses of *AP1S1* SNPs further revealed the association between AP1S1 and ESCC in different subgroups. *AP1S1* rs77387752 genotype TT was associated with increased risk in patients who never drink alcohol or smoke, while *AP1S1* rs35208462 genotype TC was associated with lower risk in the same population. The subjects who are male, younger, tobacco consumers and non-drinkers were associated with lower ESCC risk in *AP1S1* rs4729666 genotype TC. Previous study showed that ESCC is more common in males, elder population, tobacco and alcohol consuming can increase the risk of ESCC by about 3–7 times as well.[44–46] Our results seemed to some extent contradictory to the previous findings in the contributory factors to the development of ESCC. Thus, the functional relevance of *AP1S1* rs4729666 genotype TC underlying this difference demands further investigation.

Although these SNPs are not in the same functional region, the SNPs rs77387752, rs35208462 and rs4729666 were significantly linked to each other. Rs77387752 C>T is a synonymous variant in the coding sequence, while rs4729666 and rs35208462 are variant in untranslated region. In the haplotype analyses, we found that *AP1S1* rs4729666 genotype TCCACCC and *AP1S1* TGCCC were associated with increased risk of ESCC, while *AP1S1* CACCT and *AP1S1* TGCTC were associated with lower susceptibility to ESCC. Despite the underlying mechanisms are still not fully understood, our study involving 1043 cases and 1143 controls suggested that *AP1S1* may present as a promising target to study the potential mechanisms of tumorigenesis and a molecular target in targeted therapy of cancer treatment.

In the additive test, rs77387752 T>C was associated with higher risk (borderline statistically significant, P=0.067), while rs35208462 T>C and rs4729666 T>C were associated with lower risk (statistically significant, P=0.009, 0.024) with ESCC. Because the MAF of rs77387752 is only 0.058, and the haplotype with frequency less than 0.03 would not be considered in the analysis. So as for haplotype with T in rs77387752, we only identified *AP1S1* TATCC, and its OR was not statistically significant which may be due to its low frequency and borderline significance. However, in the haplotype with higher risk, *AP1S1* CACCC and *AP1S1* TGCCC, both had C in rs35208462 and rs4729666 and the haplotype with lower susceptibility, *AP1S1* CACCT and *AP1S1* TGCTC both had one T mutation in rs35208462 or rs4729666, so the result of haplotype analysis is consistence with the result of logistic regression analyses.

The mechanisms underlying how *AP1S1* variant affected the susceptibility to ESCC are still unknown so far. Alshabi et al reported that high expression of *AP1S1* is related to pathogenesis of glioblastoma and shorter survival.[47] Both hypomethylated *AP1S1* and *NEIL2* can be a novel marker of acute myeloid leukemia.[31] As a result, both the expression level and epigenetic modification of *AP1S1* are related to tumor. The rs77387752 is located in coding sequence, while the rs4729666 and rs35208462 are located in intron. The mutations at introns usually do not seriously affect protein function, however, introns play a role in the stability of RNA, efficacy of translation and the rapid induction of protein expression through post-transcriptional regulation,[48] and we speculate that the rs4729666 and rs35208462 variant work in a similar way. Then single synonymous variant can influence protein expression level through translation kinetics but not through protein conformation,[49] and this may be how the rs77387752 works. Further cell biological and mechanistic studies are warranted.

We consider this study have several strengths. First, the subjects were all from same region of China and they could be regarded as a homogenous population, therefore racial admixture were negligible. Second, this is the first study exploring the relationship between ESCC and *AP1S1* SNPs, our findings could provide novel target for cancer genesis, prevention, and treatment.

It should also be noted that there are several limitations in this study. First, subjects were all recruited from the same area in eastern China, which may not be fully representative of the general Chinese population. Second, ESCC is a complex disease with high heterogeneity. Our result only indicated the impact of single SNP of *AP1S1* in research subjects our study was in the lack of technical support to measure the copper level in serum and tumor tissue, it's still debatable to make a conclusion that these SNPs are directly risk factors for ESCC.[50] And the follow up of the therapeutic effects of chemotherapy was also deficient, so the biological functions of the SNPs of *AP1S1* in ESCC need further investigation. Third, the result was underpowered for the inherent feature of hospital-based control study and some variables may be small in subgroup analysis. Future studies among different regions and populations, with more detailed epidemiologic information and supplemented with functional analyses, are warranted to further verify our conclusions.

Conclusions

In conclusion, we found that the *AP1S1* rs35208462 and rs4729666 were associated a lower risk of ESCC, while *AP1S1* rs77387752 associated with increased susceptibility to ESCC, especially in the subjects who are younger, male, and without alcohol and smoking history.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Gupta B, Kumar N. Worldwide incidence, mortality and time trends for cancer of the oesophagus. *Eur J Cancer Prev*. 2017;26(2):107–118. doi:10.1097/CEJ.0000000000000249
3. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut*. 2015;64 (3):381–387. doi:10.1136/gutjnl-2014-308124
4. Fitzmaurice C, Dicker D, Pain A, et al.; Global Burden of Disease Cancer Collaboration. The global burden of cancer 2013. *JAMA Oncol*. 2015;1 (4):505–527. doi:10.1001/jamaoncol.2015.0735
5. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. *Lancet*. 2013;381(9864):400–412. doi:10.1016/S0140-6736(12)60643-6
6. Gammon MD, Schoenberg JB, Ahsan H, et al. Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst*. 1997;89(17):1277–1284. doi:10.1093/jnci/89.17.1277
7. Lee CH, Wu DC, Lee JM, 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(7):1188–1199. doi:10.1016/j.ejca.2007.01.039
8. 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(5):763–766. doi:10.1016/S0959-8049(05)80363-6
9. Enzinger PC, Mayer RJ. Esophageal cancer. *N Engl J Med*. 2003;349(23):2241–2252. doi:10.1056/NEJMra035010
10. Taylor PR, Qiao YL, Abnet CC, et al. Prospective study of serum vitamin E levels and esophageal and gastric cancers. *J Natl Cancer Inst*. 2003;95 (18):1414–1416. doi:10.1093/jnci/djg044

11. Abnet CC, Lai B, Qiao YL, et al. Zinc concentration in esophageal biopsy specimens measured by x-ray fluorescence and esophageal cancer risk. *J Natl Cancer Inst*. 2005;97(4):301–306. doi:10.1093/jnci/dji042

12. Businello G, Parente P, Mastracci L, et al. The pathologic and molecular landscape of esophageal squamous cell carcinogenesis. *Cancers*. 2020;12 (8):2160. doi:10.3390/cancers12082160

13. Dell'Angelica EC, Bonifacino JS. Coatopathies: genetic disorders of protein coats. *Annu Rev Cell Dev Biol*. 2019;35:131–168. doi:10.1146/annurev-cellbio-100818-125234

14. Martinelli D, Dionisi-Vici C. AP1S1 defect causing MEDNIK syndrome: a new adaptinopathy associated with defective copper metabolism. *Ann N Y Acad Sci*. 2014;1314:55–63. doi:10.1111/nyas.12426

15. Nevitt T, Ohrvik H, Thiele DJ. Charting the travels of copper in eukaryotes from yeast to mammals. *Biochim Biophys Acta*. 2012;1823 (9):1580–1593. doi:10.1016/j.bbamcr.2012.02.011

16. Brady DC, Crowe MS, Turski ML, et al. Copper is required for oncogenic BRAF signalling and tumorigenesis. *Nature*. 2014;509(7501):492–496. doi:10.1038/nature13180

17. Turski ML, Brady DC, Kim HJ, et al. A novel role for copper in Ras/mitogen-activated protein kinase signaling. *Mol Cell Biol*. 2012;32 (7):1284–1295. doi:10.1128/MCB.05722-11

18. Papa L, Manfredi G, Germain D. SOD1, an unexpected novel target for cancer therapy. *Genes Cancer*. 2014;5(1–2):15–21. doi:10.18632/genesandcancer.4

19. Ishida S, Andreux P, Poitry-Yamate C, Auwerx J, Hanahan D. Bioavailable copper modulates oxidative phosphorylation and growth of tumors. *Proc Natl Acad Sci U S A*. 2013;110(48):19507–19512. doi:10.1073/pnas.1318431110

20. Barker HE, Cox TR, Erler JT. The rationale for targeting the LOX family in cancer. *Nat Rev Cancer*. 2012;12(8):540–552. doi:10.1038/nrc3319

21. Theophanides T, Anastassopoulou J. Copper and carcinogenesis. *Crit Rev Oncol Hematol*. 2002;42(1):57–64. doi:10.1016/S1040-8428(02)00007-0

22. Yang F, Pei R, Zhang Z, et al. Copper induces oxidative stress and apoptosis through mitochondria-mediated pathway in chicken hepatocytes. *Toxicol in Vitro*. 2019;54:310–316. doi:10.1016/j.tiv.2018.10.017

23. Kang Z, Qiao N, Liu G, Chen H, Tang Z, Li Y. Copper-induced apoptosis and autophagy through oxidative stress-mediated mitochondrial dysfunction in male germ cells. *Toxicol in Vitro*. 2019;61:104639. doi:10.1016/j.tiv.2019.104639

24. Liao J, Yang F, Chen H, et al. Effects of copper on oxidative stress and autophagy in hypothalamus of broilers. *Ecotoxicol Environ Saf*. 2019;185:109710. doi:10.1016/j.ecoenv.2019.109710

25. Brady DC, Crowe MS, Greenberg DN, Counter CM. Copper chelation inhibits BRAF(V600E)-driven melanomagenesis and counters resistance to BRAF(V600E) and MEK1/2 Inhibitors. *Cancer Res*. 2017;77(22):6240–6252. doi:10.1158/0008-5472.CAN-16-1190

26. Yoshii J, Yoshiji H, Kuriyama S, et al. The copper-chelating agent, trientine, suppresses tumor development and angiogenesis in the murine hepatocellular carcinoma cells. *Int J Cancer*. 2001;94(6):768–773. doi:10.1002/ijc.1537

27. Cunzhi H, Jiexian J, Xianwen Z, Jingang G, Suling H. Classification and prognostic value of serum copper/zinc ratio in Hodgkin's disease. *Biol Trace Elem Res*. 2001;83(2):133–138. doi:10.1385/BTER:83:2:133

28. Witkowski K, Kozłowski A, Pardela M, Piecuch J, Walichiewicz P. [Level of copper in plasma and tissue of patients with esophageal and large bowel cancer]. *Wiad Lek*. 1993;46(15–16):586–588. Polish.

29. Li YQ, Chen J, Yin JY, Liu ZQ, Li XP. Gene expression and single nucleotide polymorphism of ATP7B are associated with platinum-based chemotherapy response in non-small cell lung cancer patients. *J Cancer*. 2018;9(19):3532–3539. doi:10.7150/jca.26286

30. Yang T, Chen M, Chen T, Thakur A. Expression of the copper transporters hCtr1, ATP7A and ATP7B is associated with the response to chemotherapy and survival time in patients with resected non-small cell lung cancer. *Oncol Lett*. 2015;10(4):2584–2590. doi:10.3892/ol.2015.3531

31. Mallik S, Zhao Z. Towards integrated oncogenic marker recognition through mutual information-based statistically significant feature extraction: an association rule mining based study on cancer expression and methylation profiles. *Quant Biol*. 2017;5(4):302–327. doi:10.1007/s40484-017-0119-0

32. 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(7):e101705. doi:10.1371/journal.pone.0101705

33. Yin J, Wang X, Zheng L, et al. Hsa-miR-34b/c rs4938723 T>C and hsa-miR-423 rs6505162 C>A polymorphisms are associated with the risk of esophageal cancer in a Chinese population. *PLoS One*. 2013;8(11):e80570. doi:10.1371/journal.pone.0080570

34. Chen ZJ, Zhao H, He L, et al. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet*. 2011;43(1):55–59. doi:10.1038/ng.732

35. Shi YY, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res*. 2005;15(2):97–98. doi:10.1038/sj.cr.7290272

36. Li Z, Zhang Z, He Z, et al. A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (http://analysis.bio-x.cn). *Cell Res*. 2009;19(4):519–523. doi:10.1038/cr.2009.33

37. Zanconato F, Forcato M, Battilana G, et al. Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat Cell Biol*. 2015;17(9):1218–1227. doi:10.1038/ncb3216

38. Maugeri-Saccà M, De Maria R. The Hippo pathway in normal development and cancer. *Pharmacol Ther*. 2018;186:60–72. doi:10.1016/j.pharmthera.2017.12.011

39. Klee KMC, Janecke AR, Civan HA, et al. AP1S1 missense mutations cause a congenital enteropathy via an epithelial barrier defect. *Hum Genet*. 2020;139(10):1247–1259. doi:10.1007/s00439-020-02168-w

40. Montpetit A, Côté S, Brustein E, et al. Disruption of AP1S1, causing a novel neurocutaneous syndrome, perturbs development of the skin and spinal cord. *PLoS Genet*. 2008;4(12):e1000296. doi:10.1371/journal.pgen.1000296

41. Incecik F, Bisgin A, Yılmaz M. MEDNIK syndrome with a frame shift causing mutation in AP1S1 gene and literature review of the clinical features. *Metab Brain Dis*. 2018;33(6):2065–2068. doi:10.1007/s11011-018-0313-4

42. Alonso N, Wani S, Rose L, Van't Hof RJ, Ralston SH, Albagha OME. Insertion mutation in Tnfrsf11a causes a Paget's disease-like phenotype in heterozygous mice and osteopetrosis in homozygous mice. *J Bone Miner Res*. 2021;36(7):1376–1386. doi:10.1002/jbmr.4288

43. Ferwerda B, Alonso S, Banahan K, et al. Functional and genetic evidence that the Mal/TIRAP allele variant 180L has been selected by providing protection against septic shock. *Proc Natl Acad Sci U S A*. 2009;106(25):10272–10277. doi:10.1073/pnas.0811273106

44. Kamangar F, Chow WH, Abnet CC, Dawsey SM. Environmental causes of esophageal cancer. *Gastroenterol Clin North Am*. 2009;38(1):27–57. doi:10.1016/j.gtc.2009.01.004

45. Seitz HK, Maurer B, Stickel F. Alcohol consumption and cancer of the gastrointestinal tract. *Dig Dis*. 2005;23(3–4):297–303. doi:10.1159/000090177

46. Zheng R, Zeng H, Zhang S, Chen T, Chen W. National estimates of cancer prevalence in China, 2011. *Cancer Lett*. 2016;370(1):33–38. doi:10.1016/j.canlet.2015.10.003

47. Alshabi AM, Vastrad B, Shaikh IA, Vastrad C. Identification of crucial candidate genes and pathways in glioblastoma multiform by bioinformatics analysis. *Biomolecules*. 2019;9(5):201. doi:10.3390/biom9050201

48. Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. *Hum Genet*. 2017;136(9):1043–1057. doi:10.1007/s00439-017-1791-x

49. Hunt R, Hettiarachchi G, Katneni U, et al. A single synonymous variant (c.354G>A [p.P118P]) in ADAMTS13 confers enhanced specific activity. *Int J Mol Sci*. 2019;20(22):5734. doi:10.3390/ijms20225734

50. Lin L, Lin DC. Biological significance of tumor heterogeneity in esophageal squamous cell carcinoma. *Cancers*. 2019;11(8):1156. doi:10.3390/cancers11081156