Research Paper

Validation study of susceptibility loci for esophageal squamous cell carcinoma identified by GWAS in a Han Chinese subgroup from Eastern China

Kai-Lai Wang[1*], Xiang-Liu Chen[1*], Lan Lei[1*], Pei Li[2*], Lian-Lian Hong[1], Xian-Chong Huang[1], Wei-Min Mao[3,4], Kenichi Mukaisho[5], Zhi-Qiang Ling[1,4✉]

1. Zhejiang Cancer Institute, Zhejiang Cancer Hospital, No.1 Banshan East Rd., Gongshu District, Hangzhou 310022, P.R.China.
2. Department of Pathophysiology, School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, 450052, China.
3. Department of Thoracic Tumor Surgery, Zhejiang Cancer Hospital, No.1 Banshan East Rd., Gongshu District, Hangzhou 310022, P.R.China.
4. Zhejiang Key Laboratory of Diagnosis & Treatment Technology on Thoracic Oncology (Lung and Esophagus), Hangzhou 310022, China.
5. Department of Pathology, Division of Molecular Diagnostic Pathology, Shiga University of Medical Science, Otsu, Shiga, Japan.

*Kai-Lai Wang, Xiang-Liu Chen, Lan Lei and Pei Li wish it to be known that, in their opinion, the first four authors should be regarded as joint-first authors.

✉ Corresponding author: Zhi-Qiang Ling, Professor, MD.,PhD, Zhejiang Cancer Institute, Zhejiang Cancer Hospital, Zhejiang Cancer Center. No.1 Banshan East Road, Gongshu District, Hangzhou 310022, P. R. China. Tel: +86-571-88122423; Fax: +86-571-88122423; E-mail: lingzq@zjcc.org.cn or lingzq@hotmail.com

Abstract

Esophageal squamous cell carcinoma (ESCC) occurs at a relatively high frequency in China and is one of the most prevalent cancers in the world. Genome-wide association studies (GWAS) have identified 24 single-nucleotide polymorphisms (SNPs) that could be associated with ESCC in Chinese patients. This retrospective study aimed to validate the association between these 24 SNPs and ESCC in a Han Chinese subgroup from East China. A total of 2280 and 1900 patients with ESCC (case group) and non-esophageal cancer (control group) were included from a single center. Genotyping of the 24 polymorphisms was performed using the Sequenom MassARRAY system. Unconditional logistic regression analyses were conducted for every polymorphism. It was found that rs12188136 (P=0.027, OR=1.158, 95% CI=1.016-1.319 for AG/AA) was associated with ESCC. Binary logistic regression analyses revealed a significant negative association of rs875339 in *RORA* (P=0.014, OR=0.762, 95% CI=0.613-0.947 for TT/CC). Under the dominant model, rs6854472 was slightly associated with ESCC risk (P=0.048, OR=1.192, 95% CI=1.002-1.418). Under the recessive model, a significant negative association was observed for rs875339 (P=0.010, OR=0.758, 95% CI=0.615-0.935). In a word, this large-scale replication study validated that rs12188136 and rs6854472 are associated with ESCC in a Han Chinese subgroup from Eastern China, and that rs875339 is negative associated with ESCC.

Key words: esophagi al squamous cell carcinoma (ESCC) , genome-wide association study (GWAS), single nucleotide polymorphism (SNP), MassARRAY system, Han Chinese population

Introduction

Esophageal cancer is among the most incident malignant tumors worldwide [1] and a serious threat to human health and quality of life [2]. Esophageal squamous cell carcinoma (ESCC) is one of the two main sub-types of esophageal cancer, and ESCC is more common than esophageal adenocarcinoma in the developing world, especially in China [3]. The prognosis of ESCC is poor despite advances in treatment, with 5-year overall survival rate ranging from 15% to 25% [4,5].

Accumulating evidence has demonstrated that genetic factors [6-10], family history of ESCC [11-13], lifestyle habits [14-16], environmental factors [17-23], and HPV infection [24] play important roles in the

development of ESCC. Significant interactions were found between HPV serological status and genetic loci, increasing the risk of ESCC [25,26]. Other risk factors such as exposure to polycyclic aromatic hydrocarbons (PAHs), high-temperature foods, diets, oral health and microbial communities, but they require further research. Esophageal carcinogenesis is the result of the interaction among heredity, environment and living habits [27-29].

In recent years, genome-wide association studies (GWAS) have confirmed the contribution of gene variations to ESCC [30-35]. Six large-scale GWAS of Chinese populations have focused on identifying genetic susceptibility loci for ESCC [31-35]. The earliest ESCC GWAS analysis using 2115 ESCC cases and 3302 controls in a Chinese population revealed that PLCE1 carried cancer susceptibility [31]. Wang *et al*. identified two new genome-wide significant loci for ESCC: PLCE1 at 10q23 and C20orf54 at 20p13 [32]. Seven loci on chromosomes 5q11, 6p21, 10q23, 12q24 and 21q22 were associated with the risk of ESCC [33]. In another GWAS in a Chinese ESCC population, Wu *et al*. [34] identified nine new ESCC susceptibility loci: seven (on chromosomes 4q23, 16q12.1, 17q21, 22q12, 3q27, 17p13 and 18p11) had a significant marginal effect on the risk of ESCC and two (on 2q22 and 13q33) had a significant association but only when considering the gene-alcohol interaction. Wu *et al*. identified rs1050631 in SLC39A6 as being associated with the survival of ESCC patients [35].

Whether those 24 SNPs found by the five GWAS confer an increased risk of ESCC in various Han Chinese populations has not yet been validated. Therefore, we conducted a case-control study to validate the associations of those 24 SNPs with the risk of ESCC in a Han Chinese subgroup from Eastern China.

Material and Methods

Study population

This was a retrospective study. We included 2280 consecutive ESCC subjects and 1900 non-ESCC subjects (control group). The diagnosis of ESCC was confirmed by histopathology or cytology by at least two local pathologists. Histological examination was performed according to the World Health Organization (WHO) criteria [36]. The exclusion criteria for both groups were: 1) psychiatric disorder; 2) any other primary cancer; or 3) a family history of cancer. This study consisted of two ESCC sets: (a) 1900 patients with primary ESCC, and (b) 380 patients with second ESCC. The patients were recruited between January 2012 and December 2014 at Zhejiang Cancer Hospital. Demographic characteristics of the subjects (including

gender, age, histological types of esophageal cancer, smoking and drinking status) were obtained from the medical records. Non-ESCC individuals (n=1900) were recruited as control subjects during a routine health check-up (physical examination) at the same hospital during the same time period. The two groups were matched based on the frequency of age and sex. In the present study, all participants were ethnic Han Chinese that lived within the Zhejiang Province of Eastern China.

SNP selection

We selected the 24 top SNPs (rs4478858, rs10881372, rs10801638, rs10173378, rs888103, rs3815501, rs6717108, rs10934685, rs6768588, rs9824873, rs6854472, rs12188136, rs2294693, rs9364414, rs7916519, rs11225815, rs10895458, rs4578395, rs11059556, rs2025245, rs9584006, rs347940, rs875339, and rs12922317) from the reports focusing on ESCC susceptibility loci identified by five GWAS projects in Han Chinese (PubMed search) [31-35].

SNP genotyping assays

Venous blood (2 mL) was sampled in citrate glass tubes and kept at -40°C. Leukocyte total genomic DNA was extracted from 1 mL of peripheral blood using the Whole Blood DNA Extraction Kit (QIAamp® DNA Blood Mini Kit), according to the manufacturer's instructions. The extracted genomic DNA was dissolved in 0.1× TE buffer (10 mMTris and 1 mM EDTA, pH 8.0) to 0.4-0.6 mg/mL and stored at -20°C.

The SNPs were determined using iPLEX chemistry on a matrix-assisted laser desorption/ionization time-of-flight mass spectrometer (MALDI-TOF-MS, MassARRAY system, Sequenom, Inc.), as previously published [37]. PCR reactions (5 μL each) were carried out in 384-well plates using 10 ng of genomic DNA, 0.5 units of Taq polymerase (HotStar-Taq, Qiagen), 500 μmol of each of the four deoxy-nucleotides triphosphate (dNTP), and 100 nmol of each primer. An ABI-9700 thermocycler was used with the following program: 1) 15 min at 94°C; 2) 45 cycles of 20 s at 94°C, 30 s at 56°C, and 60 s at 72°C. The reaction products were separated on 2.0% agarose. After PCR, 0.3 units of shrimp alkaline phosphatase was added and incubated at 37°C for 20 min followed by inactivation for 5 min at 85°C. The concentration of the extension primers was adjusted to optimize the signal-to-noise ratio. The iPLEX Gold Kits (Sequenom, Inc.) was used to prepare the samples with 0.2 μL (100 μmol) of termination mix, 0.05 units of DNA polymerase (Sequenom, Inc.), and 625 to 1250 nmol/L extension primers. The iPLEX reaction was performed using the following program:

1) initial denaturation for 30 s at 94°C; 2) 5 s at 94°C and five cycles of 5 s at 52°C and 5 s at 80°C; 3) 40 annealing and extension cycles; 4) 5 s at 94°C; 5) five cycles of 5 s at 52°C and 5 s at 80°C; and 6) 72°C for 3 min and the sample The products were analyzed by MALDI-TOF-MS. The samples were desalted using 6 mg of resin and transferred to a 384-well Spectro-CHIP (Sequenom, Inc.). The mass spectra were acquired and analyzed using the MassARRAYTyper 4.0 Software (Sequenom, Inc.). Controls were performed without template DNA. All laboratory technicians were unaware of patient status.

Statistical analyses

Values were expressed as means ± standard deviation (SD) or numbers. Continuous variables were analyzed using the unpaired Student's *t*-test. Differences in frequencies of the alleles and genotypes between case group and control group were evaluated using the χ^2-test. Genotype distribution and allele frequencies were compared using the chi-square test. The chi-square test was also used to examine the Hardy-Weinberg Equilibrium (HWE) in the control group (*P*-value of <0.05 was considered to be statistically significant). Akaike's information criteria were used to select the most parsimonious genetic model for each SNP [38]. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by unconditional logistic regression analysis. All analyses were conducted with Stata statistical package (version 10.0; Stata Corp LP, College Station, TX, USA). The *P* value of allele difference was conducted with chi-square test between esophageal cancer and control group. *P*-value < 0.05 was considered statistically significant.

Results

Characteristics of the subjects

The demographic characteristics of the subjects are shown in Table 1. There were no differences in age (57.0±8.8 vs. 56.4±9.3 years) or gender (male, 63.1% vs. 64.5%) between the two groups (both *P* > 0.05).

Table 1. Demographic characteristics of ESCC cases and controls used in the study

Study	N	Age, mean (s.d.)	Sex, male (%)
Cases	2280	57.0 (8.8)	64.6
First ESCC	1900	57.0 (9.4)	64.5
Second ESCC	380	56.8 (9.0)	64.7
Controls	1900	56.4 (9.3)	64.5

Individual SNP association analysis

The genomic characteristics of 24 SNPs are given in Table 2. There was no deviation from the Hardy-Weinberg equilibrium in the control group (all *P* > 0.01). In the single-locus analyses, the allelic frequencies of rs10173378: A>G (0.241 vs. 0.221, *P* = 0.0409) and rs6854472: G>T (0.072 vs. 0.084, *P* = 0.0477) were slightly different between the ESCC and control group, but 100,000 permutations showed that there were no significant differences between the two groups. The genotype distributions of the 24 SNPs in the two groups are summarized in Table 3. The distribution of the rs12188136 (47.4% vs. 50.2%, *P* = 0.0493) and rs875339 (49.4% vs. 48.4%, *P* = 0.0341) genotypes showed significant differences between the cases and controls.

Logistic regression analyses revealed that in the codominant-effect model, the ESCC risk was associated with rs12188136 (*P* = 0.027, OR = 1.158, 95% CI = 1.016-1.319 for AG/AA). Binary logistic regression analyses revealed a slight negative association of rs10895458 (*P* = 0.044, OR = 0.547, 95% CI = 0.304-0.983 for CC/AA) and a significant negative association of rs875339 (*P* = 0.014, OR = 0.762, 95% CI = 0.613-0.947 for TT/CC), but because of the rarity of the homozygous mutant genotype (<3%), the results were invalid for rs10895458. In addition, marginal esophageal cancer risk was found for rs6854472 (*P* = 0.056, OR = 1.187, 95% CI = 0.995-1.417 for GT/GG) (Table 3).

Using the dominant model, significant ESCC risk was observed for rs6854472 (*P* = 0.048, OR = 1.192, 95% CI = 1.002-1.418). Using the recessive model, a significant negative association was observed for rs875339 (*P* = 0.010, OR = 0.758, 95% CI = 0.615-0.935) (Table 4).

Discussion

ESCC is one of the most prevalent cancers worldwide and occurs at a relatively high frequency in China. Some recent genome-wide association studies have identified 24 single-nucleotide polymorphisms that may be associated with ESCC. This study aimed to validate the association between these 24 polymorphisms and ESCC in a Han subgroup from Eastern China. The results suggest that rs12188136 and rs6854472 are associated with ESCC in this Han Chinese subgroup, and that rs875339 is negative associated with ESCC.

This study was a large-scale study in Han Chinese patients from Eastern China that describes the association between ESCC and 24 genome-wide SNPs. Besides rs12188136 and rs6854472 localizing in intergenic areas, *RORA* could play a role in the development of ESCC [31-35]. Abnet *et al.* [31] conducted the first large-scale genome-wide association studies for ESCC using 2115 ESCC cases and 3302 controls in Chinese, and identified PLCE1 at

10q23 for ESCC susceptibility. Then, Wang *et al*. [32] performed a GWAS of ESCC by genotyping 1077 individuals with ESCC and 1733 control subjects of Han Chinese descent, and found that PLCE1 and C20orf54 play important roles for ESCC carcinogenesis. Wu *et al*. [33] performed a GWAS on 2031 ESCC individuals and 2044 controls of Chinese descent, and evaluated promising associations in an additional 6276 cases and 6165 controls from different areas of China. They identified five chromosomal regions (5q11, 6p21, 10q23, 12q24 and 21q22) that carried seven susceptibility loci for ESCC in the Chinese population, of which three (5q11, 6p21 and 21q22) were newly discovered [33]. Wu *et al*. [34] reported a multistage GWAS of ESCC in 10,123 ESCC cases and 10,664 controls. This GWAS identified nine new susceptibility loci for ESCC, of which seven (4q23, 16q12.1, 17q21, 22q12, 3q27, 17p13 and 18p11) had a significant marginal effect and two of which (2q22 and 13q33) had a significant association in the gene-alcohol interaction only [34]. Among 5337 Chinese with ESCC and 5787 controls (replication in 9654 Chinese with ESCC and 10,058 controls), Wu *et al*. [34] showed that rs7447927at 5q31.2 and rs1642764 at 17p13.1 were associated with ESCC susceptibility [34]. Furthermore, Hu *et al*. [39] showed that rs2274223 was associated with reduced PLCE1 expression and increased risk of ESCC. Another replication study by Wang *et al*. [40] showed that the ADH1B-ADH1C-ADH7 axis was modulated by the rs1042026, rs17033, rs1614972, rs1789903 and rs17028973 SNPs. In the present study, the identified polymorphisms matched those found by the previous studies, and included rs2294693 in 6p21.1, rs11059556 in 12q24, rs6854472 in 4q22, rs12922317 in 16p13.12, and rs9824873 in 3q28. The discrepancies among studies regarding the identified loci can be due to the genetic diversity among different regions of China and of the world. Additional studies are necessary to better understand the risk of ESCC.

Table 2. Information about 24 validated SNPs.

Gene: locus and OMIM No.[a]	SNP_ID	Chromosome No.	Chromosome Position [b]	Reference allele	Effect allele	MAF[c]			Pg	P value for HWE[h] test	Genotyping call Rate (%)[i]
						NCBI[d]	control[e]	EC[f]			
SERINC2: 1p35.1 OMIN: 614549	rs4478858	1	31411078	G	A	0.200	0.196	0.213	0.0592	0.551	96.75
1p13	rs10881372	1	106210655	C	T	0.163	0.196	0.187	0.2864	0.224	97.37
1q31	rs10801638	1	198002090	C	T	0.349	0.302	0.304	0.8751	0.795	97.13
2p22	rs10173378	2	43119650	A	G	0.198	0.241	0.221	0.0409	0.600	97.18
LYPD6: 2q23.2 OMIN: 613359	rs888103	2	149370922	C	T	0.128	0.115	0.119	0.6138	0.464	97.32
BZW1: 2q33 OMIN: N.A	rs3815501	2	200821399	G	A	0.488	0.464	0.467	0.8099	0.448	97.15
2q36	rs6717108	2	224696318	C	T	0.444	0.444	0.449	0.6156	0.265	96.82
UMPS: 3q21.2 OMIN: 613891	rs10934685	3	124747673	C	T	0.389	0.341	0.335	0.5626	0.677	96.65
ITGB5: 3q21.2 OMIN: 147561	rs6768588	3	124768488	A	G	0.244	0.278	0.282	0.7215	0.475	96.60
3q28	rs9824873	3	183583986	T	C	0.291	0.329	0.325	0.7073	0.928	96.17
4q22	rs6854472	4	89513521	G	T	0.085	0.072	0.084	0.0477	0.871	97.75
5q35	rs12188136	5	174407635	A	G	0.256	0.296	0.305	0.3698	0.222	96.77
UNC5CL: 6p21.1 OMIN: N.A	rs2294693	6	41037763	T	C	0.267	0.253	0.245	0.4418	0.074	97.13
6q27	rs9364414	6	168171267	G	A	0.360	0.376	0.387	0.3088	0.439	96.79
10p12	rs7916519	10	23177805	G	A	0.140	0.230	0.234	0.6363	0.929	97.01
DYNC2H1: 11q22.3 OMIN: 603297	rs11225815	11	103469085	T	C	0.233	0.255	0.248	0.4884	0.606	96.56
11q22	rs10895458	11	103547356	A	C	0.133	0.113	0.101	0.0938	0.266	97.30
OPCML: 11q25 OMIN: 600632	rs4578395	11	133242868	T	C	0.105	0.091	0.088	0.7086	0.252	96.82
12q24.3	rs11059556	12	128161518	C	T	0.279	0.336	0.338	0.8632	0.826	96.41
13q13	rs2025245	13	37529440	G	A	0.354	0.381	0.364	0.1169	0.401	97.22
GPC5: 13q31.3 OMIN: 602446	rs9584006	13	92249673	T	G	0.372	0.422	0.406	0.1536	0.746	96.39
FMN1: 15q13.3 OMIN: 136535	rs347940	15	32885469	A	G	0.442	0.357	0.357	0.9613	0.553	95.96
RORA: 15q22.2 OMIN: 600825	rs875339	15	60803856	C	T	0.314	0.313	0.295	0.0881	0.024	97.01
SNX29: 16p13.13-p13.12 OMIN: N.A	rs12922317	16	11983775	G	A	0.256	0.318	0.300	0.0916	0.369	96.82

a. OMIM, Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/Omim); **b.** SNP position in the NCBI dbSNP Build 38 database (http://www.ncbi.nlm.nih.gov/SNP); **c.** MAF, minor allele frequency, representing the frequency of effect allele; **d.** MAF for Chinese in the NCBI dbSNPs database; **e.** MAF for control group; **f.** MAF for esophageal cancer group; **g.** P value, which was conducted with χ^2 test, for difference in allele distributions between esophageal cancer and control group; **h.** HWE, Hardy–Weinberg equilibrium in control group; **i**. The percentage of successful genotype calls.

Table 3. Genotype frequencies of 24 validated SNPs among cases and control and their associations with esophageal cancer risk under co-dominant genetic model.

Gene	SNP ID	Genotype	Case		Control		P (2 df)[a]	Logistic regression		P$_{trend}$
			No.	Frequency (%)	No.	Frequency (%)		OR (95%CI)	P[b]	
SERINC2	rs4478858	GG	1380	62.22	1185	64.90	0.1732	1.000 (reference)		0.061
		GA	732	33.00	567	31.05		1.109 (0.969-1.268)	0.133	
		AA	106	4.78	74	4.05		1.230 (0.905-1.672)	0.186	
	rs10881372	CC	1481	66.00	1172	64.18	0.4701	1.000 (reference)		0.282
		CT	688	30.66	592	32.42		0.920 (0.804-1.052)	0.221	
		TT	75	3.34	62	3.40		0.957 (0.678-1.352)	0.804	
	rs10801638	CC	1096	49.04	891	48.82	0.7977	1.000 (reference)		0.876
		CT	920	41.16	765	41.92		0.978 (0.858-1.114)	0.734	
		TT	219	9.80	169	9.26		1.053 (0.846-1.312)	0.641	
	rs10173378	AA	1341	60.16	1053	57.45	0.1040	1.000 (reference)		0.039
		AG	789	35.40	678	36.99		0.914 (0.802-1.041)	0.176	
		GG	99	4.44	102	5.56		0.762 (0.571-1.017)	0.065	
LYPD6	rs888103	CC	1743	77.95	1432	78.17	0.3918	1.000 (reference)		0.615
		CT	456	20.39	379	20.69		0.988 (0.848-1.152)	0.882	
		TT	37	1.65	21	1.15		1.448 (0.844-2.484)	0.179	
BZW1	rs3815501	GG	633	28.33	516	28.24	0.8648	1.000 (reference)		0.809
		GA	1115	49.91	925	50.63		0.983 (0.850-1.136)	0.813	
		AA	486	21.75	386	21.13		1.026 (0.860-1.225)	0.773	
	rs6717108	CC	689	30.94	575	31.59	0.8831	1.000 (reference)		0.620
		CT	1075	48.27	875	48.08		1.025 (0.889-1.182)	0.731	
		TT	463	20.79	370	20.33		1.044 (0.876-1.245)	0.629	
UMPS	rs10934685	CC	972	43.82	787	43.19	0.8142	1.000 (reference)		0.560
		CT	1006	45.36	827	45.39		0.985 (0.864-1.123)	0.821	
		TT	240	10.82	208	11.42		0.934 (0.759-1.150)	0.522	
ITGB5	rs6768588	AA	1139	51.19	939	51.79	0.9298	1.000 (reference)		0.719
		AG	919	41.30	740	40.82		1.024 (0.899-1.166)	0.722	
		GG	167	7.51	134	7.39		1.027 (0.806-1.310)	0.827	
	rs9824873	TT	1014	45.84	816	45.13	0.9025	1.000 (reference)		0.708
		TC	960	43.40	796	44.03		0.971 (0.851-1.107)	0.656	
		CC	238	10.76	196	10.84		0.977 (0.792-1.206)	0.830	
	rs6854472	GG	1884	83.92	1586	86.15	0.1370	1.000 (reference)		0.047
		GT	347	15.46	246	13.36		1.187 (0.995-1.417)	0.056	
		TT	14	0.62	9	0.49		1.310 (0.565-3.033)	0.529	
	rs12188136	AA	1050	47.36	917	50.16	0.0493	1.000 (reference)		0.367
		AG	981	44.25	740	40.48		**1.158 (1.016-1.319)**	**0.027**	
		GG	186	8.39	171	9.35		0.950 (0.758-1.191)	0.656	
UNC5CL	rs2294693	TT	1268	56.78	1035	56.65	0.1972	1.000 (reference)		0.445
		TC	835	37.39	661	36.18		1.031 (0.905-1.175)	0.647	
		CC	130	5.82	131	7.17		0.810 (0.627-1.047)	0.107	
	rs9364414	GG	818	36.83	718	39.34	0.1831	1.000 (reference)		0.307
		GA	1086	48.90	841	46.08		1.133 (0.991-1.297)	0.068	
		AA	317	14.27	266	14.58		1.046 (0.864-1.267)	0.645	
	rs7916519	GG	1306	58.62	1085	59.39	0.8840	1.000 (reference)		0.636
		GA	801	35.95	645	35.30		1.032 (0.905-1.176)	0.641	
		AA	121	5.43	97	5.31		1.036 (0.784-1.370)	0.802	
DYNC2H1	rs11225815	TT	1265	56.73	999	55.32	0.6251	1.000 (reference)		0.488
		TC	824	36.95	694	38.43		0.938 (0.823-1.069)	0.334	
		CC	141	6.32	113	6.26		0.985 (0.759-1.279)	0.912	
	rs10895458	AA	1800	80.65	1450	79.02	0.0916	1.000 (reference)		0.094
		AC	413	18.50	357	19.46		0.932 (0.796-1.091)	0.381	
		CC	19	0.85	28	1.53		**0.547 (0.304-0.983)**	**0.044**	
OPCML	rs4578395	TT	1845	82.88	1501	82.43	0.9295	1.000 (reference)		0.705
		TC	368	16.53	309	16.97		0.969 (0.821-1.144)	0.709	
		CC	13	0.58	11	0.60		0.961 (0.430-2.152)	0.924	
	rs11059556	CC	961	43.29	799	44.14	0.6764	1.000 (reference)		0.863
		CT	1016	45.77	804	44.42		1.051 (0.921-1.199)	0.462	
		TT	243	10.95	207	11.44		0.976 (0.793-1.201)	0.819	
	rs2025245	GG	907	40.55	708	38.75	0.2705	1.000 (reference)		0.119
		GA	1030	46.04	845	46.25		0.951 (0.832-1.088)	0.467	
		AA	300	13.41	274	15.00		0.855 (0.706-1.035)	0.107	
GPC5	rs9584006	TT	780	35.15	609	33.65	0.3271	1.000 (reference)		0.154
		TG	1077	48.54	876	48.40		0.960 (0.836-1.103)	0.563	
		GG	362	16.31	325	17.96		0.870 (0.724-1.045)	0.136	
FMN1	rs347940	AA	911	41.00	745	41.64	0.6201	1.000 (reference)		0.961
		AG	1037	46.67	810	45.28		1.047 (0.916-1.196)	0.500	
		GG	274	12.33	234	13.08		0.958 (0.784-1.169)	0.670	
RORA	rs875339	CC	1104	49.44	881	48.35	0.0341	1.000 (reference)		0.091
		CT	939	42.05	742	40.72		1.010 (0.886-1.151)	0.883	
		TT	190	8.51	199	10.92		**0.762 (0.613-0.947)**	**0.014**	

Gene	SNP ID	Genotype	Case		Control		P (2 df)[a]	Logistic regression		P$_{trend}$
			No.	Frequency (%)	No.	Frequency (%)		OR (95%CI)	P[b]	
SNX29	rs12922317	GG	1090	49.05	841	46.08	0.1695	1.000 (reference)		0.091
		GA	929	41.81	808	44.27		0.887 (0.779-1.011)	0.072	
		AA	203	9.14	176	9.64		0.890 (0.713-1.110)	0.301	

a. Global P values [2 degrees of freedom (df)]: genotype frequencies in esophageal cancer and control group were compared using a χ² test with 2 df. **b.** P values from unconditional logistic regression analyses.

Table 4. Association analysis of 24 validated SNPs under dominant and recessive genetic model.

Gene	SNP ID	Genetic model	Case	Control	Logistic regression[a]	
					OR (95%CI)	P[c]
SERINC2	rs4478858	(GA+AA) vs. GG	838/1380	641/1185	1.123 (0.987-1.277)	0.079
		AA vs. (GG+GA)	106/2112	74/1752	1.188 (0.877-1.610)	0.265
	rs10881372	(CT+TT) vs. CC	763/1481	654/1172	0.923 (0.811-1.051)	0.227
		TT vs. (CC+CT)	75/2169	62/1764	0.984 (0.699-1.385)	0.925
	rs10801638	(CT+TT) vs. CC	1139/1096	934/891	0.991 (0.876-1.122)	0.891
		TT vs. (CC+CT)	219/2016	169/1656	1.064 (0.862-1.314)	0.562
	rs10173378	(AG+GG) vs. AA	888/1341	780/1053	0.894 (0.788-1.014)	0.080
		GG vs. (AA+AG)	99/2130	102/1731	0.789 (0.594-1.048)	0.101
LYPD6	rs888103	(CT+TT) vs. CC	493/1743	400/1432	1.013 (0.872-1.176)	0.870
		TT vs. (CC+CT)	37/2199	21/1811	1.451 (0.846-2.488)	0.176
BZW1	rs3815501	(GA+AA) vs. GG	1601/633	1311/516	0.995 (0.868-1.142)	0.948
		AA vs. (GG+GA)	486/1748	386/1441	1.038 (0.893-1.207)	0.628
	rs6717108	(CT+TT) vs. CC	1538/689	1245/575	1.031 (0.902-1.178)	0.655
		TT vs. (CC+CT)	463/1764	370/1450	1.029 (0.882-1.199)	0.718
UMPS	rs10934685	(CT+TT) vs. CC	1246/972	1035/787	0.975 (0.860-1.105)	0.688
		TT vs. (CC+CT)	240/1978	208/1614	0.942 (0.773-1.146)	0.549
ITGB5	rs6768588	(AG+GG) vs. AA	1086/1139	874/939	1.024 (0.905-1.160)	0.704
		GG vs. (AA+AG)	167/2058	134/1679	1.017 (0.803-1.288)	0.890
	rs9824873	(TC+CC) vs. TT	1198/1014	992/816	0.972 (0.858-1.101)	0.654
		CC vs. (TT+TC)	238/1974	196/1612	0.992 (0.812-1.211)	0.934
	rs6854472	(GT+TT) vs. GG	361/1884	255/1586	**1.192 (1.002-1.418)**	**0.048**
		TT vs. (GG+GT)	14/2231	9/1832	1.277 (0.552-2.958)	0.568
	rs12188136	(AG+GG) vs. AA	1167/1050	911/917	1.119 (0.988-1.266)	0.076
		GG vs. (AA+AG)	186/2031	171/1657	0.887 (0.714-1.103)	0.282
UNC5CL	rs2294693	(TC+CC) vs. TT	965/1268	792/1035	0.995 (0.878-1.127)	0.932
		CC vs. (TT+TC)	130/2103	131/1696	0.800 (0.623-1.029)	0.082
	rs9364414	(GA+AA) vs. GG	1403/818	1107/718	1.112 (0.979-1.264)	0.101
		AA vs. (GG+GA)	317/1904	266/1559	0.976 (0.818-1.164)	0.785
	rs7916519	(GA+AA) vs. GG	922/1306	742/1085	1.032 (0.910-1.171)	0.620
		AA vs. (GG+GA)	121/2107	97/1730	1.024 (0.778-1.348)	0.864
DYNC2H1	rs11225815	(TC+CC) vs. TT	965/1265	807/999	0.944 (0.833-1.070)	0.369
		CC vs. (TT+TC)	141/2089	113/1693	1.011 (0.783-1.306)	0.932
	rs10895458	(AC+CC) vs. AA	432/1800	385/1450	0.904 (0.775-1.054)	0.198
		CC vs. (AA+AC)	19/2213	28/1807	0.554 (0.308-0.995)	0.048
OPCML	rs4578395	(TC+CC) vs. TT	381/1845	320/1501	0.969 (0.823-1.141)	0.702
		CC vs. (TT+TC)	13/2213	11/1810	0.967 (0.432-2.163)	0.934
	rs11059556	(CT+TT) vs. CC	1259/961	1011/799	1.035 (0.914-1.173)	0.586
		TT vs. (CC+CT)	243/1977	207/1603	0.952 (0.782-1.159)	0.623
	rs2025245	(GA+AA) vs. GG	1330/907	1119/708	0.928 (0.818-1.053)	0.245
		AA vs. (GG+GA)	300/1937	274/1553	0.878 (0.736-1.048)	0.149
GPC5	rs9584006	(TG+GG) vs. TT	1439/780	1201/609	0.935 (0.821-1.066)	0.318
		GG vs. (TT+TG)	362/1857	325/1485	0.891 (0.756-1.050)	0.168
FMN1	rs347940	(AG+GG) vs. AA	1311/911	1044/745	1.027 (0.905-1.165)	0.680
		GG vs. (AA+AG)	274/1948	234/1555	0.935 (0.775-1.127)	0.479
RORA	rs875339	(CT+TT) vs. CC	1129/1104	941/881	0.957 (0.846-1.084)	0.491
		TT vs. (CC+CT)	190/2043	199/1623	**0.758 (0.615-0.935)**	**0.010**
SNX29	rs12922317	(GA+AA) vs. GG	1132/1090	984/841	0.888 (0.784-1.005)	0.060
SERINC2		AA vs. (GG+GA)	203/2019	176/1649	0.942 (0.762-1.165)	0.581

a. P values from unconditional logistic regression analyses.

The present study is not without limitations. Statistical correction was used to adjust for multiple testing for a specific gene, but this is controversial. The Bonferroni correction and Bayesian techniques are frequently used, but they are problematic when correcting multiple comparisons [41] and such corrections might not be needed when different associations are of interest on a purely one-at-a-time basis [42,43]. Secondly, our study included patients with first ESCC and second ESCC. First ESCC is more relevant to genetic factors than second ESCC. Thirdly, although our study suggested that some loci may be involved in the prevalence of acquired ESCC, only selected SNPs based on the literature were examined and they might not be enough to describe the entire genetic variation of Han Chinese. Finally, this was a

retrospective study and data about lifestyle habits (especially smoking and drinking) were not available or reliable for all patients, preventing subgroup and interaction analyses. Beyond the association studies, the literature is currently limited by the lack of mechanistic studies about the involvement of these SNPs in the development of ESCC and the present study was not designed to determine those mechanisms. Additional studies will have to be carried out on this issue.

Conclusion

This large-scale replication study showed that rs12188136 and rs6854472 are associated with ESCC in a Han Chinese subgroup from Eastern China, and that rs875339 is negative associated with ESCC. This study underlines the genetic complexity of ESCC development.

Acknowledgements

Author contributions

Wang KL, Chen XL, Lei L, Li P and Ling ZQ contributed to the design, execution, and analysis of this paper. Wang KL, Chen XL, Lei L, Li P and Ling ZQ drafted the manuscript. Hong LL and Huang XC provided some help for data analysis. All the authors (including Mao WM) were involved in the critical revision of the manuscript.

Ethics statement

The study was performed in accordance with the Declaration of Helsinki and approved by the Ethics Committees of Zhejiang Cancer Hospital. Written informed consent was obtained for the recruitment of each participant. Each participant was then interviewed to collect detailed information on demographic characteristics.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Gupta B, Kumar N. Worldwide incidence, mortality and time trends for cancer of the oesophagus. Eur J Cancer Prev. 2017; 26:107-118.
2. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010; 127:2893-917.
3. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. Gastroenterology. 2018; 154:360-373.
4. Zeng H, Chen W, Zheng R, et al. Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. Lancet Glob Health. 2018; 6:e555-e567.
5. Chen W, Sun K, Zheng R, et al. Cancer incidence and mortality in China, 2014. Chin J Cancer Res. 2018; 30:1-12.
6. Zhou L, Fu G, Wei J, et al. The identification of two regulatory ESCC susceptibility genetic variants in the TERT-CLPTM1L loci. Oncotarget. 2016; 7: 5495-506.
7. Li K, Yin X, Yang H, et al. Association of the genetic polymorphisms in XRCC6 and XRCC5 with the risk of ESCC in a high-incidence region of North China. Tumori. 2015; 101:24-9.
8. Zhang H, Li H, Yu H. Analysis of the role of rs2031920 and rs3813867 polymorphisms within the cytochrome P450 2E1 gene in the risk of squamous cell carcinoma. Cancer Cell Int. 2018; 18:67.
9. Moghadam AR, Mehramiz M, Entezari M, et al. A genetic polymorphism in the CYP1B1 gene in patients with squamous cell carcinoma of the esophagus: an Iranian Mashhad cohort study recruited over 10 years. Pharmacogenomics. 2018; 19:539-546.
10. Yin J, Tang W, Long T, et al. 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. Oncotarget. 2017; 8: 110153-110165.
11. Chen T, Cheng H, Chen X, et al. Family history of esophageal cancer increases the risk of esophageal squamous cell carcinoma. Sci Rep. 2015; 5:16038.
12. Bhat GA, Shah IA, Rafiq R, et al. Family history of cancer and the risk of squamous cell carcinoma of oesophagus: a case-control study in Kashmir, India. Br J Cancer. 2015; 113:524-32.
13. Jia N, Wen X, Zhang N, et al. Younger age of onset and multiple primary lesions associated with esophageal squamous cell carcinoma cases with a positive family history of the cancer suggests genetic predisposition. Chin Med J (Engl). 2014; 127: 2779-83.
14. Chang J, Huang Y, Wei L, et al. Risk prediction of esophageal squamous-cell carcinoma with common genetic variants and lifestyle factors in Chinese population. Carcinogenesis 2013;34:1782-6.
15. Jessri M, Rashidkhani B, Hajizadeh B, et al. Adherence to Mediterranean-style dietary pattern and risk of esophageal squamous cell carcinoma: a case-control study in Iran. J Am Coll Nutr. 2012; 31:338-51.
16. Wu IC, Wu CC, Lu CY, et al. Substance use (alcohol, areca nut and cigarette) is associated with poor prognosis of esophageal squamous cell carcinoma. PloS one 2013; 8:e55834.
17. Taccioli C, Chen H, Jiang Y, et al. Dietary zinc deficiency fuels esophageal cancer development by inducing a distinct inflammatory signature. Oncogene 2012; 31:4550-8.
18. Guo W, Zhao YP, Jiang YG, et al. Restoring the metabolic disturbance of zinc: may not only contribute to the prevention of esophageal squamous cell cancer. Med Hypotheses. 2008; 71:957-9.
19. Tan W, Miao X, Wang L, et al. Significant increase in risk of gastroesophageal cancer is associated with interaction between promoter polymorphisms in thymidylate synthase and serum folate status. Carcinogenesis. 2005; 26:1430-5.
20. Stolzenberg-Solomon RZ, Qiao YL, Abnet CC, et al. Esophageal and gastric cardia cancer risk and folate- and vitamin B(12)-related polymorphisms in Linxian, China. Cancer Epidemiol Biomarkers Prev. 2003; 12:1222-6.
21. Oka M, Yamamoto K, Takahashi M, et al. Relationship between serum levels of interleukin 6, various disease parameters and malnutrition in patients with esophageal squamous cell carcinoma. Cancer Res. 1996; 56:2776-80.
22. Chen CH, Lu HI, Wang YM, et al. Areca nut is associated with younger age of diagnosis, poor chemoradiotherapy response, and shorter overall survival in esophageal squamous cell carcinoma. PloS one. 2017; 12:e0172752.
23. De Stefani E, Deneo-Pellegrini H, Ronco AL, et al. Meat consumption, cooking methods, mutagens, and risk of squamous cell carcinoma of the esophagus: a case-control study in Uruguay. Nutr Cancer. 2012; 64:294-9.
24. Cao F, Han H, Zhang F, et al. HPV infection in esophageal squamous cell carcinoma and its relationship to the prognosis of patients in northern China. ScientificWorldJournal. 2014; 2014:804738.

25. Yang J, Liu B, Li W, et al. Association of p53 and MDM2 polymorphisms with risk of human papillomavirus (HPV)-related esophageal squamous cell carcinoma (ESCC). Cancer Epidemiol. 2013; 37:629-33.

26. Yang J, Wu H, Wei S, et al. HPV seropositivity joints with susceptibility loci identified in GWASs at apoptosis associated genes to increase the risk of Esophageal Squamous Cell Carcinoma (ESCC). BMC cancer. 2014; 14:501.

27. Tang WR, Chen ZJ, Lin K, et al. Development of esophageal cancer in Chaoshan region, China: association with environmental, genetic and cultural factors. Int J Hyg Environ Health. 2015; 218:12-8.

28. Chen J, Kwong DL, Cao T, et al. Esophageal squamous cell carcinoma (ESCC): advance in genomics and molecular genetics. Dis Esophagus. 2015; 28:84-9.

29. Gholipour M, Islami F, Roshandel G, et al. Esophageal Cancer in Golestan Province, Iran: A Review of Genetic Susceptibility and Environmental Risk Factors. Middle East J Dig Dis. 2016; 8:249-266.

30. Hu N, Wang C, Hu Y, et al. Genome-wide association study in esophageal cancer using GeneChip mapping 10K array. Cancer Res. 2005; 65:2542-6.

31. Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. Nat Genet. 2010; 42:764-7.

32. Wang LD, Zhou FY, Li XM, et al. Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. Nat Genet. 2010; 42:759-63.

33. Wu C, Hu Z, He Z, et al. Genome-wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in Chinese populations. Nat Genet. 2011; 43:679-84.

34. Wu C, Kraft P, Zhai K, et al. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. Nat Genet. 2012; 44:1090-7.

35. Wu C, Li D, Jia W, et al. Genome-wide association study identifies common variants in SLC39A6 associated with length of survival in esophageal squamous-cell carcinoma. Nat Genet. 2013; 45:632-8.

36. Lordick F, Mariette C, Haustermans K, et al. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016; 27(suppl 5):v50-v57.

37. Zhao LQ, Wen ZJ, Wei Y, et al. Polymorphisms of renin-angiotensin-aldosterone system gene in chinese han patients with nonfamilial atrial fibrillation. PloS one 2015; 10:e0117489.

38. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19:716-23.

39. Hu H, Yang J, Sun Y, et al. Putatively functional PLCE1 variants and susceptibility to esophageal squamous cell carcinoma (ESCC): a case-control study in eastern Chinese populations. Ann Surg Oncol. 2012; 19:2403-10.

40. Wang J, Wei J, Xu X, et al. Replication study of ESCC susceptibility genetic polymorphisms locating in the ADH1B-ADH1C-ADH7 cluster identified by GWAS. PloS one. 2014; 9:e94096.

41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Statist Soc B. 1995; 57:289-300.

42. Perneger TV. What's wrong with Bonferroni adjustments. Bmj. 1998; 316:1236-8.

43. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990; 1:43-6.