



# Gastric cancer may share genetic predisposition with esophageal squamous cell carcinoma in Chinese populations

Linhua Yao<sup>1,2,3,4</sup> · Fei Yu<sup>1,2</sup> · Yingying Mao<sup>5</sup> · Tianpei Wang<sup>1,2</sup> · Qi Qi<sup>1,2</sup> · Hui Ding<sup>1,2</sup> · Jinchun Wang<sup>1,2</sup> · Hongxia Ma<sup>1,2</sup> · Juncheng Dai<sup>1,2</sup> · Guoxin Zhang<sup>3</sup> · Guangfu Jin<sup>1,2</sup>

Received: 8 May 2018 / Revised: 15 July 2018 / Accepted: 31 July 2018 / Published online: 10 September 2018  
© The Author(s) under exclusive licence to The Japan Society of Human Genetics 2018

## Abstract

Many features are shared between esophageal cancer (EC) and gastric cancer (GC). This study aimed to explore whether known EC susceptibility loci are also important in the development of GC. A total of 21 genetic variants associated with EC in genome-wide association studies were evaluated with association of GC risk in 2631 cases and 4373 controls of Chinese ancestry. Single variant and weighted genetic scores (WGS) for esophageal squamous cell carcinoma (ESCC), esophageal adenocarcinoma (EAC), and overall EC were analyzed with GC risk, respectively. Genetic variants of rs2274223 in *PLCE1* at 10q23.33 (per G allele: odds ratio (OR) = 1.26, 95% confidence interval (CI): 1.16–1.38,  $P = 6.51 \times 10^{-8}$ ), rs10052657 in *PDE4D* at 5q11.2 (per C allele: OR = 1.12, 95% CI: 1.01–1.25,  $P = 3.28 \times 10^{-2}$ ) and rs671 in *ALDH2* at 12q24.12 (per A-allele: OR = 0.83, 95% CI: 0.75–0.91,  $P = 1.14 \times 10^{-4}$ ) were significantly associated with GC risk. The combined effect of those three variants had stronger influence on GC risk (OR = 1.31, 95% CI: 1.19–1.44,  $P = 2.34 \times 10^{-8}$ ). High WGS of ESCC was also associated with increased risk of GC ( $P = 5.52 \times 10^{-4}$  as a continuous variable) (trend test  $P = 2.71 \times 10^{-4}$ ), whereas no statistically significant associations were observed between the WGS of EAC and GC risk ( $P = 0.66$  as a continuous variable) (trend test  $P = 0.70$ ). ESCC rather than EAC may share genetic susceptibility with GC. Genetic variants at 10q23.33, 5q11.2, and 12q24.12 may be useful as biomarkers to identify individuals with high risk for both ESCC and GC.

## Introduction

Esophageal cancer (EC) and gastric cancer (GC) are two common gastrointestinal cancers worldwide, with 456,000 new cases and 400,000 estimated deaths per year for EC, and

951,000 new cases and 723,000 estimated deaths per year for GC, respectively [1]. The stomach is connected to the esophagus through gastroesophageal junction, which is also known as cardia. Based on the anatomical location, GC can be classified into two types: true gastric (non-cardia) and gastroesophageal junction cancers (cardia) [2]. The majority of GC is gastric adenocarcinoma, whereas EC consists of two histopathological types, esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC) [3]. ESCC is predominant around the world especially in China, whereas

Authors share co-first authorship: Linhua Yao, Fei Yu, Yingying Mao

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1038/s10038-018-0501-4>) contains supplementary material, which is available to authorized users.

✉ Guangfu Jin  
guangfujin@njmu.edu.cn  
✉ Guoxin Zhang  
guoxinz@njmu.edu.cn

<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China  
<sup>2</sup> Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine, Nanjing Medical University, Nanjing, China

<sup>3</sup> Department of Gastroenterology, First Affiliated Hospital of Nanjing Medical University, Nanjing, China  
<sup>4</sup> Department of Gastroenterology, The First Affiliated Hospital of Huzhou Teachers College 158 Guang Chuang Hou Lu, Huzhou 313000 Zhejiang, China  
<sup>5</sup> Department of Epidemiology and Biostatistics, School of Public Health, Zhejiang Chinese Medical University, Hangzhou 310053 Zhejiang, China

EAC subtype is a major type in the United States, Australia, the United Kingdom, and other European countries [4].

Both environmental and genetic factors contribute to the development of EC and GC. There are similarities and differences in the risk factors for ESCC and EAC. Tobacco use was associated with increased risk of both ESCC and EAC [5–7]. Alcohol consumption is a specific risk factor for ESCC [8, 9], whereas gastroesophageal reflux disease, obesity, and Barrett's esophagus were associated with increased susceptibility of EAC [9]. The known environmental risk factors for GC are *Helicobacter pylori* infection, smoking, obesity, low intake of fresh fruits and vegetables, and high consumption of salted foods [2]. Genome-wide association studies (GWAS) have been conducted to explore genetic variants influencing the susceptibility of EC and GC over the past few decades [10–12]. A missense mutation located in *PLCE1* named rs2274223 was found to be associated with risk of ESCC and gastric cardia cancer [10].

In consideration of the close anatomical location and similarities among risk factors between EC and GC, we hypothesized that the genetic basis of developing EC and GC might have something in common. To test whether single EC risk variant or cumulative genetic risk score computed using established EC risk loci were also associated with GC risk, we utilized risk loci reported in the published EC GWAS and tested whether they were associated with GC in our large case-control studies.

## Materials and methods

### Study populations

Participants of the current study were from three published GC GWAS. The Nanjing GWAS (565 cases and 1162 controls) and the Beijing GWAS (468 cases and 1123 controls) were based on two independent case-control studies, which were reported previously [12]. For the National Cancer Institute (NCI) GWAS (1625 cases and 2100 controls), subjects were recruited from Shanxi and Linxian [10]. All cases in the Nanjing and Beijing GWAS were diagnosed with non-cardia GC, whereas cases in the NCI GWAS contained both cardiac and non-cardia GC. In total, 2631 GC cases and 4373 controls were included in the current study. Basic demographic information of the participants was shown in Supplementary table 1.

### Genetic variants selection

We searched the GWAS catalog (<https://www.genome.gov/gwastudies/>, last accessed on 25 July 2017) for genetic variants associated with EC risk. Besides, we searched PubMed database (<https://www.ncbi.nlm.nih.gov/pubmed/>) for recently published EC GWAS. The reported EC risk loci

were filtered using the following criteria: (1) the reported significance level of the association reaching  $5.00 \times 10^{-8}$ ; (2) the minor allele frequency (MAF) of variants not  $< 1\%$  in the Chinese population (1000 Genomes phase III); (3) for variants in linkage disequilibrium (LD) defined as  $r^2 > 0.1$ , we selected the variant with the lowest  $P$  value. Finally, nine GWAS (eight studies for ESCC, one study for EAC) were included in our current study [10, 11, 13–19]. 42 single nucleotide polymorphisms (SNP) reached the predefined significance level of association, but four had the MAF of  $< 1\%$  in the Chinese population. After excluding SNPs in LD, 21 SNPs (18 SNPs for ESCC, 3 SNPs for EAC) were included in the final statistical analysis. Detailed information about the eligible EC risk loci are shown in Supplementary table 2.

### Imputation and quality control

After basic quality-control procedures performed in GWAS, we excluded SNPs with call rate  $< 95\%$ , MAF  $< 0.01$ , and Hardy-Weinberg equilibrium  $P$  value  $< 1.00 \times 10^{-6}$ . Then, we performed imputation for the Nanjing (Affymetrix 6.0), the Beijing (Affymetrix 6.0) and the NCI (Illumina 660 W) GWAS separately using software SHAPEIT [20] and IMPUTE2 [21]. We used all populations from the 1000 Genomes Project Phase III as the reference set. After imputation, we further excluded SNPs with poor imputation quality (info score  $< 0.3$ ) and repeated the quality control procedures for SNPs mentioned above. Among the 21 selected SNPs, we did not get genotype information of rs76014404 and rs8030672 from the Nanjing/Beijing or the NCI GWAS. Therefore, we used two SNPs (rs2143771 and rs116760846) in complete LD with these SNPs in the following analyses (Supplementary table 3).

### Calculation of weighted genetic scores

Weighted genetic score (WGS) was calculated to evaluate the cumulative effect of esophagus cancer risk loci on GC risk. We calculated two independent WGS for ESCC and EAC, as there was high heterogeneity in the genetic background between these two subtypes. We also combined the two subtypes to measure the WGS for overall EC. For each individual, WGS was calculated by multiplying the number of risk alleles by the EC-associated beta ( $\beta_j$ ), which was derived from published studies. For rs2274223, which was reported in more than one study, we estimated its effect on EC based on meta-analysis. To calculate WGS for the  $i$ th subject, the following formula was used:

$$WGS_i = \sum_1^j x_{ij}\beta_j$$

In this formula,  $x_{ij}$  is the number of risk alleles for the  $j$ -th variant in the  $i$ -th subject ( $x_{ij}=0, 1, \text{ or } 2$ ) and  $\beta_j$  is the coefficient or weight for the  $j$ -th variant (calculated by ln-transformed of odds ratios (ORs) from published studies).

### Differential expression analysis

Expression data (normalized expectation-maximization read counts) were downloaded from the Cancer Genome Atlas database, which consisted of 87 EAC tissues, 10 EAC paired normal tissues, 85 ESCC tissues, three ESCC-paired normal tissues, 413 GC tissues, and 32 GC paired normal tissues, respectively. Expression data were log<sub>2</sub> transformed to correspond to normal distribution. Paired  $t$  test (10 EAC tissues vs. 10 EAC paired normal tissues, 3 ESCC tissues vs. 3 ESCC paired normal tissues, 32 GC tissues vs. 32 GC paired normal tissues) and two-sample  $t$  test (87 EAC tissues vs. 10 EAC paired normal tissues, 85 ESCC tissues vs. 3 ESCC paired normal tissues, 413 GC tissues vs. 32 GC paired normal tissues) were used to evaluate differential expression among tumor and normal tissues.

### Statistical analysis

Genetic association analysis was conducted by using logistic regression models. When dealing with association between single locus and GC risk, we assumed an additive genetic model in logistic regression. For GC risk-associated variants, we estimated the cumulative effect based on the number of risk alleles. We included the WGS in the logistic regression model both as a continuous variable and categorical variable. For the Nanjing and the Beijing GWAS, we adjusted for age, sex, smoking, drinking status, and principal component analysis (PCA) for population stratification, and for the NCI GWAS, we adjusted for age, sex, and PCA in the regression models. Meta-analysis was used to combine results from the three GWAS, and Cochran's  $Q$  was used for heterogeneity test. Fixed-effect model was applied to assume the combined effect, whereas random effect model was repeated if  $I^2$  (calculated by  $100\% \times (Q - (n-1))/Q$ ) was  $> 75\%$ . Differential expression analysis was performed based on two-sample  $t$  test or paired  $t$  test. Analyses were performed with Stata version 11 or R version 3.2.1, unless otherwise noted.

## Results

### Association between single variant and GC risk

As shown in Table 1, among the 21 known genetic variants associated with EC risk, three were significantly associated with GC risk. Consistent with the previous report [10], the

G allele of rs2274223 at 10q23.33 (reported gene: *PLCE1*) was associated with increased GC risk (OR = 1.26, 95% confidence interval (CI): 1.16–1.38,  $P = 6.51 \times 10^{-8}$ ). Considering that *PLCE1* locus was identified as a common locus between ESCC and GC in the previous GWAS using NCI samples, we reanalyzed the association between rs2274223 and GC risk by excluding samples from NCI and found the direction of the association was consistent but not significant (per G allele: OR = 1.07, 95% CI: 0.94–1.23,  $P = 0.31$ ). In addition, rs10052657 at 5q11.2 (reported gene: *PDE4D*) and rs671 at 12q24.12 (reported gene: *ALDH2*) were also associated with GC risk (OR = 1.12, 95% CI: 1.01–1.25,  $P = 3.28 \times 10^{-2}$ ; OR = 0.83, 95% CI: 0.75–0.91,  $P = 1.14 \times 10^{-4}$ , respectively). Nevertheless, there were obvious heterogeneities among three studies for all the three discovered SNPs, and they were insignificant when random effect model was used in meta-analysis. We did not find significant associations with GC risk for the remaining 18 variants.

We further examined the cumulative effect of these three variants (rs2274223, rs10052657, and rs671) on GC risk (Table 2). We found a strong tendency of increased GC risk with greater numbers of risk alleles (OR = 1.31, 95% CI: 1.19–1.44,  $P = 2.34 \times 10^{-8}$ ).

### Association of EC WGS with GC risk

Because the observed effect of rs671 on ESCC and GC was in opposite directions, we derived the WGS based on the reported effect size of the 20 variants (excluding rs671) from the original EC study (Supplementary table 1) and evaluated the association of EC WGS and risk of GC (Table 3 and Table 4). We found that the EC WGS was significantly associated with increased risk of GC (OR = 1.15, 95% CI: 1.06–1.25,  $P = 1.20 \times 10^{-3}$  for continuous WGS and OR = 1.08, 95% CI: 1.03–1.13,  $P = 9.11 \times 10^{-4}$  for trend test for WGS categories). This association was mainly restricted to ESCC (OR = 1.16, 95% CI: 1.07–1.27,  $P = 5.52 \times 10^{-4}$  for continuous WGS and OR = 1.09, 95% CI: 1.04–1.14,  $P = 2.71 \times 10^{-4}$  for trend test for WGS categories) but not to EAC (OR = 1.02, 95% CI: 0.92–1.13,  $P = 0.66$  for continuous WGS and OR = 0.99, 95% CI: 0.95–1.04,  $P = 0.70$  for trend test for WGS categories).

### Differential expression analysis of candidate genes

We further analyzed whether the expression levels of the genes associated with those three variants (rs2274223, rs10052657, and rs671) were altered in cancer tissues compared with normal tissues of esophagus and stomach (Supplementary figure 1). We found that *PDE4D* and *ALDH2* were downregulated in both ESCC and GC tissues as compared with normal tissues. However, we did not

**Table 1** Associations of 21 known EC risk variants with GC risk

Cytoband	SNP	EH/Ref <sup>a</sup>	Gene <sup>b</sup>	Cancer	Nanjing <sup>c</sup>			Beijing <sup>c</sup>			NCT <sup>c</sup>			Meta <sup>c</sup>						
					Stat <sup>c</sup>	Info <sup>d</sup>	OR(95% CI)	P	Stat <sup>c</sup>	Info <sup>d</sup>	OR(95% CI)	P	Stat <sup>c</sup>	Info <sup>d</sup>	OR(95% CI)	P	$I^2$ (%)			
2p24.1	rs7255	C/T	<i>GDF7</i> , <i>C2orf43</i>	EAC	I	0.85	0.90	0.23	I	0.85	1.01	0.94	G	I	1.03	0.60	0.99	0.88	0.19	0
3q27.1	rs9823696	A/G	<i>HTR3C</i>	EAC	I	0.94	0.95	0.55	I	0.95	1.00	0.98	I	0.98	0.91	0.09	0.93	0.11	0.38	0
2q33.1	rs13016963	A/G	<i>RUNX1</i>	ESCC	I	0.95	0.95	0.55	I	0.89	0.89	0.24	G	I	1.02	0.76	0.98	0.58	0.22	0
3q27.3	rs2239612	A/G	<i>ST6GALI</i>	ESCC	G	1.13	1.13	0.20	G	1.09	1.09	0.46	G	I	1.05	0.43	1.07	0.13	0.50	0
4q21.23	rs1494961	C/T	<i>HELO</i>	ESCC	G	1.16	1.16	0.06	G	0.87	0.87	0.12	G	I	0.90	0.03	0.95	0.16	2.69 × 10 <sup>-3</sup>	77.79
4q23	rs1042026	T/C	<i>ADH1B</i>	ESCC	G	1.03	1.03	0.78	G	1.19	1.19	0.08	G	I	0.98	0.68	1.02	0.59	0.08	33.31
4q23	rs1789903	G/C	<i>ADH1C</i>	ESCC	G	0.75	0.75	3.85 × 10 <sup>-2</sup>	G	0.91	0.91	0.52	I	I	0.98	0.85	0.92	0.18	0.09	31.96
4q23	rs971074	T/C	<i>ADH7</i>	ESCC	I	0.68	0.98	0.90	I	0.69	1.16	0.32	I	I	1.03	0.64	1.04	0.49	0.39	0
5q11.2	rs10052657	C/A	<i>PDE4D</i>	ESCC	I	0.76	1.43	6.00 × 10 <sup>-3</sup>	I	0.76	1.03	0.82	G	I	1.07	0.28	1.12	3.28 × 10 <sup>-2</sup>	4.02 × 10 <sup>-2</sup>	52.51
5q31.2	rs7447927	G/C	<i>TMEM173</i>	ESCC	I	0.96	0.96	0.61	I	0.96	1.21	3.62 × 10 <sup>-2</sup>	I	0.97	0.92	0.09	0.98	0.49	8.18 × 10 <sup>-3</sup>	71.4
6p21.1	rs10484761	C/T	<i>UNC5CL</i>	EAC	I	0.99	0.96	0.74	I	0.99	1.12	0.43	I	0.99	0.94	0.42	0.97	0.64	0.28	0
6p21.32	rs35597309	A/G	<i>HLA-DRBI</i> , <i>HLA-DQA1</i>	ESCC	I	0.96	1.08	0.64	I	0.96	0.94	0.71	I	0.99	1.01	0.88	1.01	0.87	0.55	0
6q11.1	rs76014404	T/TAAACA	<i>KHDRBS2</i>	ESCC	I	1.21	1.21	0.21	I	1.18	1.18	0.32	I	0.99	1.01	0.95	1.08	0.29	0.23	0
10q23.33	rs2274223	G/A	<i>PLCE1</i>	ESCC	G	1.05	1.05	0.59	G	1.10	1.10	0.35	G	I	1.41	9.05 × 10 <sup>-10</sup>	1.26	6.51 × 10 <sup>-8</sup>	2.01 × 10 <sup>-3</sup>	79.04
12q24.12	rs671	A/G	<i>ALDH2</i>	ESCC	I	0.98	0.56	2.76 × 10 <sup>-9</sup>	I	0.98	0.91	0.44	G	I	0.95	0.46	0.83	1.14 × 10 <sup>-4</sup>	3.35 × 10 <sup>-6</sup>	90.74
12q24.13	rs4767364	G/A	<i>ALDH2</i>	ESCC	I	1.18	1.18	0.27	I	1.16	1.16	0.34	G	I	1.00	0.98	1.06	0.39	0.25	0
15q23	rs8030672	A/T	<i>ANP32A</i>	ESCC	I	0.67	0.82	0.59	I	0.63	0.83	0.64	I	0.77	1.22	0.23	1.10	0.50	0.22	0

Table 1 (continued)

Cytoband	SNP	Eff/Ref <sup>a</sup>	Gene <sup>b</sup>	Cancer	Nanjing <sup>c</sup>			Beijing <sup>c</sup>			Meta <sup>f</sup>									
					Sta <sup>c</sup>	Info <sup>d</sup>	OR(95% CI)	P	Sta <sup>c</sup>	Info <sup>d</sup>	OR(95% CI)	P	OR(95% CI)	P	I <sup>2</sup> (%)					
16q12.1	rs4785204	T/C	HEATR3	ESCC	I	I	0.94 (0.80,1.11)	0.49	I	I	1.13 (0.93,1.37)	0.21	I	I	1.04 (0.94,1.16)	0.42	1.03 (0.95,1.12)	0.42	0.15	2.73
17p13.1	rs1642764	T/C	ATP1B2, TP53	ESCC	I	I	0.87 (0.83,1.11)	0.59	I	I	0.93 (0.78,1.10)	0.39	I	I	0.92 (0.92,1.11)	0.83	0.98 (0.92,1.06)	0.65	0.35	0
21q22.12	rs2014300	A/G	RUNX1	ESCC	G	I	0.91 (0.72,1.15)	0.42	G	I	0.72 (0.57,0.91)	6.36 × 10 <sup>-3</sup>	I	I	0.94 (0.89,1.18)	0.72	0.93 (0.84,1.03)	0.19	1.10 × 10 <sup>-2</sup>	69.06
22q12.1	rs4822983	T/C	CHEK2	ESCC	I	I	0.77 (0.86,1.20)	0.83	I	I	0.77 (0.96,1.40)	0.13	I	I	0.99 (0.88,1.12)	0.91	1.03 (0.95,1.12)	0.48	0.17	0

ESCC esophageal squamous cell carcinoma, EAC esophageal adenocarcinoma, GC gastric cancer, I imputed, G genotyped, OR(95% CI) odds ratio (95% confidence interval)

<sup>a</sup>Effect allele/reference allele

<sup>b</sup>Genes extracted from the original GWAS are associated with reported EC risk loci

<sup>c</sup>Information about whether the genetic variant was imputed or genotyped in three GWAS. For rs8030672 and rs76014404, relevant information of SNPs in complete linkage disequilibrium was listed

<sup>d</sup>Info score represents imputation quality

<sup>e</sup>Results from multivariable logistic regression analyses. Nanjing: adjusted for sex, age, smoking, drinking, and PCA; Beijing: adjusted for sex, age, smoking, drinking, and PCA; NCI: adjusted for age, sex, and PCA

<sup>f</sup> Results from meta-analysis using fixed-effect models

**Table 2** Combined analysis of association between three identified genetic variants and risk of GC

Number of risk allele <sup>a</sup>	Nanjing <sup>b</sup>		Beijing <sup>b</sup>		NCI <sup>b</sup>		Meta <sup>c</sup>	
	OR(95% CI)	P	OR(95% CI)	P	OR(95% CI)	P	OR(95% CI)	P
0–2	1.00		1.00		1.00		1.00	
3–4	2.56(1.72,3.91)	6.32 × 10 <sup>-6</sup>	1.11(0.70,1.78)	0.67	1.11(0.87,1.44)	0.40	1.35(1.11,1.64)	2.80 × 10 <sup>-3</sup>
5–6	3.15(2.02,5.03)	7.82 × 10 <sup>-7</sup>	1.21(0.73,2.02)	0.47	1.53(1.16,2.02)	2.53 × 10 <sup>-3</sup>	1.72(1.39,2.14)	6.72 × 10 <sup>-7</sup>
Trend	1.56(1.28,1.90)	9.28 × 10 <sup>-6</sup>	1.09(0.87,1.38)	0.44	1.29(1.14,1.46)	5.13 × 10 <sup>-5</sup>	1.31(1.19,1.44)	2.34 × 10 <sup>-8</sup>

OR (95% CI) odds ratio (95% confidence interval)

<sup>a</sup>The number of risk allele were calculated using three genetic variants (rs2274223, rs10052657, and rs671) for each individual

<sup>b</sup>Results from multiple logistic regression. Nanjing: adjusted for sex, age, smoking, drinking, and PCA; Beijing: adjusted for sex, age, smoking, drinking, and PCA; NCI: adjusted for age, sex, and PCA

<sup>c</sup>Results from meta-analysis using fixed-effect model

**Table 3** Association between WGS (as a continuous variable) and risk of GC

	ESCC(17 SNPs) <sup>a</sup>		EAC(3 SNPs)		EC(20 SNPs) <sup>a</sup>	
	OR(95% CI)	P	OR(95% CI)	P	OR(95% CI)	P
Nanjing <sup>b</sup>	1.12(0.93,1.35)	0.24	1.37(0.69,2.73)	0.37	1.13(0.95,1.36)	0.17
Beijing <sup>b</sup>	1.22(0.99,1.51)	0.06	1.19(0.53,2.67)	0.67	1.22(1.00,1.50)	0.05
NCI <sup>b</sup>	1.16(1.04,1.30)	6.59 × 10 <sup>-3</sup>	0.73(0.47,1.13)	0.16	1.13(1.02,1.26)	2.15 × 10 <sup>-2</sup>
meta <sup>c</sup>	1.16(1.07,1.27)	5.52 × 10 <sup>-4</sup>	1.02(0.92,1.13)	0.66	1.15(1.06,1.25)	1.20 × 10 <sup>-3</sup>

ESCC esophageal squamous cell carcinoma, EAC esophageal adenocarcinoma, EC esophageal cancer, GC gastric cancer, WGS weighted genetic score, OR (95% CI) odds ratio (95% confidence interval)

<sup>a</sup>Rs671 was excluded from the WGS

<sup>b</sup>Results from multivariable logistic regression analyses. Nanjing: adjusted for sex, age, smoking, drinking, and PCA; Beijing: adjusted for sex, age, smoking, drinking, and PCA; NCI: adjusted for age, sex, and PCA

<sup>c</sup>Results from meta-analysis using fixed-effect model

observe differential expression for *PLCE1* in either ESCC or GC tissues.

## Discussion

In the current study, we investigated whether the known EC risk loci were associated with GC risk using 2631 GC cases and 4373 controls of Chinese ancestry. We found that the G allele of rs2274223, C allele of rs10052657, and G allele of rs671 were associated with increased risk of GC, and higher WGS of ESCC was associated with increased risk of GC.

*PLCE1* is located on chromosome 10q23, which encodes a phospholipase enzyme that catalyzes the hydrolysis of phosphatidylinositol-4,5-bisphosphate to generate two second messengers: inositol 1,4,5-triphosphate and diacylglycerol [22]. In addition, it also interacts with small monomeric GTPases of the Ras and Rho families and heterotrimeric G proteins [23]. Thus, *PLCE1* regulates various processes affecting cell growth, survival, differentiation, gene expression, and oncogenesis. Several studies found that the missense variation of rs2274223 in *PLCE1* was significantly associated with ESCC [24] and gastric cardia

cancer [25], but not EAC [26], which is consistent with our findings.

Rs10052657 was another locus significantly associated with GC. However, no additional study reported the association between rs10052657 and the risk of GC and EC. Rs10052657 is located in intron 5 of *PDE4D*, a gene that hydrolyzes the second messenger cAMP (cyclic adenosine monophosphate) and acts as a signal transduction molecule in multiple cell types. Previous studies have identified genetic variants in *PDE4D* were associated with risk of several cancers including breast cancer [27]. *PDE4D* was also found to be a diver gene participated in the development of cancer, and involved in cancer progression by accelerating proliferation [28–30]. *PDE4D* was overexpressed in prostate cancer [30, 31], whereas some *PDE4D* isoforms were downregulated [32, 33]. In our study, we observed down-regulation of *PDE4D* in both ESCC and GC tissues. These findings support the biological plausibility that genetic variants in *PDE4D* may confer altered risk to ESCC and GC, whereas the potential mechanism may involve different *PDE4D* isoforms as reported in studies of prostate cancer.

Unlike rs10052657, whose C allele increased risk of the above two cancers, A-allele of rs671 promoted ESCC but

**Table 4** Association between WGS category and risk of GC

	ESCC(17 SNPs) <sup>a†</sup>			EAC(3 SNPs) <sup>a†</sup>			EC(20 SNPs) <sup>a†</sup>		
	WGS <sup>b</sup>	OR(95% CI)	P	WGS <sup>b</sup>	OR(95% CI)	P	WGS <sup>b</sup>	OR(95% CI)	P
Nanjing <sup>c</sup>	Q1(<2.506)	1.18(0.87,1.61)	0.28	Q1(<0.306)	1.56(1.17,2.09)	2.50 × 10 <sup>-3</sup>	Q1(<2.828)	1.27(0.94,1.72)	0.12
	Q2(2.506,2.883)	1.56(1.16,2.10)	3.76 × 10 <sup>-3</sup>	Q2(0.306,0.307)	1.34(0.95,1.89)	0.09	Q2(2.828,3.230)	1.55(1.15,2.10)	4.31 × 10 <sup>-3</sup>
	Q3(2.883,3.290)	1.25(0.92,1.71)	0.15	Q3(0.307,0.464)	1.20(0.91,1.61)	0.20	Q3(3.230,3.630)	1.34(0.99,1.81)	0.06
	Q4(≥3.290)	1.10(1.00,1.21)	0.06	Q4(≥0.464)	1.02(0.94,1.12)	0.60	Q4(≥3.630)	1.11(1.01,1.22)	3.19 × 10 <sup>-2</sup>
	Trend			Trend			Trend		
Beijing <sup>c</sup>	Q1(<2.506)	1.09(0.77,1.54)	0.63	Q1(<0.306)	0.84(0.60,1.18)	0.31	Q1(<2.828)	1.37(0.98,1.94)	0.07
	Q2(2.506,2.883)	1.22(0.87,1.72)	0.26	Q2(0.306,0.307)	1.01(0.69,1.48)	0.96	Q2(2.828,3.230)	1.22(0.86,1.73)	0.27
	Q3(2.883,3.290)	1.27(0.90,1.81)	0.17	Q3(0.307,0.464)	1.09(0.79,1.51)	0.60	Q3(3.230,3.630)	1.38(0.97,1.98)	0.07
	Q4(≥3.290)	1.09(0.98,1.21)	0.13	Q4(≥.464)	1.05(0.95,1.17)	0.34	Q4(≥3.630)	1.08(0.97,1.21)	0.15
	Trend			Trend			Trend		
NCF <sup>c</sup>	Q1(<2.506)	1		Q1(<0.306)	0.95(0.79,1.14)	0.60	Q1(<2.828)	1.28(1.06,1.54)	1.04 × 10 <sup>-2</sup>
	Q2(2.506,2.883)	1.22(1.01,1.47)	3.95 × 10 <sup>-2</sup>	Q2(0.306,0.307)	0.89(0.71,1.10)	0.28	Q2(2.828,3.230)	1.23(1.02,1.49)	2.88 × 10 <sup>-2</sup>
	Q3(2.883,3.290)	1.26(1.04,1.52)	1.56 × 10 <sup>-2</sup>	Q3(0.307,0.464)	0.89(0.75,1.07)	0.21	Q3(3.230,3.630)	1.27(1.05,1.53)	1.22 × 10 <sup>-2</sup>
	Q4(≥3.290)	1.30(1.08,1.56)	5.30 × 10 <sup>-3</sup>	Q4(≥0.464)	0.96(0.91,1.02)	0.17	Q4(≥3.630)	1.07(1.01,1.13)	2.85 × 10 <sup>-2</sup>
	Trend			Trend			Trend		
Meta <sup>d</sup>	Q1(<2.506)	1.19(1.03,1.37)	2.11 × 10 <sup>-2</sup>	Q1(<0.306)	1.05(0.91,1.20)	0.52	Q1(<2.828)	1.29(1.12,1.49)	5.13 × 10 <sup>-4</sup>
	Q2(2.506,2.883)	1.32(1.14,1.52)	1.89 × 10 <sup>-4</sup>	Q2(0.306,0.307)	1.00(0.85,1.18)	1.00	Q2(2.828,3.230)	1.30(1.12,1.50)	4.23 × 10 <sup>-4</sup>
	Q3(2.883,3.290)	1.28(1.11,1.48)	6.25 × 10 <sup>-4</sup>	Q3(0.307,0.464)	0.99(0.86,1.14)	0.89	Q3(3.230,3.630)	1.30(1.13,1.50)	3.43 × 10 <sup>-4</sup>
	Q4(≥3.290)	1.09(1.04,1.14)	2.71 × 10 <sup>-4</sup>	Q4(≥0.464)	0.99(0.95,1.04)	0.70	Q4(≥3.630)	1.08(1.03,1.13)	9.11 × 10 <sup>-4</sup>
	Trend			Trend			Trend		

ESCC esophageal squamous cell carcinoma, EAC esophageal adenocarcinoma, EC esophageal cancer, GC gastric cancer, WGS weighted genetic score, OR (95% CI) odds ratio (95% confidence interval)

<sup>a</sup>Rs671 was excluded from the WGS

<sup>b</sup>WGS was divided into four categories based on quartiles of WGS in controls from three GWAS

<sup>c</sup>Results from multivariable logistic regression analyses. Nanjing: adjusted for sex, age, smoking, drinking, and PCA; Beijing: adjusted for sex, age, smoking, drinking, and PCA; NCI: adjusted for age, sex, and PCA

protected from GC. The SNP rs671 is located in the twelfth exon of *ALDH2* at 12q24.12. *ALDH2* belongs to the aldehyde dehydrogenase family and participates in pathway of alcohol metabolism. Several studies have shown that *ALDH2* was associated with susceptibility to cancers including GC [34, 35], head and neck cancer [36], and colorectal cancer [37]. Rs671 has been reported to be associated with ESCC [38, 39], and overall EC [40], but the conclusions were inconsistent. Rs671 was also reported to influence GC risk, though their findings were opposite to ours [41, 42]. Alcohol consumption and rs671 were considered simultaneously when evaluating their associations with GC risk in those studies. Recently, one study reported that rs671 may not increase gastric cardia adenocarcinoma (GCA) susceptibility in Chinese Han populations, but the proportion of *ALDH2* mutated allele carriers in GCA high-incidence areas was lower than that in low-incidence areas [43]. The above evidence suggested that the mutated allele may have a potential role in decreasing GC risk. In addition, Subjects with mutated *ALDH2* exhibited a lower level of alcohol consumption than wild *ALDH2* carriers [44]. Alcohol consumption is an established risk factor for cancers [45, 46] and healthy lifestyle like controlling alcohol intake is benefit for keeping cancers away. In conclusion, effect of rs671 on GC might depend on consumption of alcohol and A-allele may turn from a protective role to hazardous role in GC if subjects intake high level of alcohol.

From the analysis based on WGS, we observed a significant association between GC risk and WGS of ESCC rather than EAC. It suggested that ESCC and GC may share common genetic background. ESCC is the main histological type of EC in Asian countries, whereas the incidence of EAC now exceeds ESCC in European and American countries. There are great differences between the two histological types of EC in pathophysiology and pathogenesis [4]. However, most ESCC associated loci were discovered based on participants from Asian. In consideration of the prevalence of ESCC and GC in China, and the well-known shared genetic variant rs2274223 between the two cancers, there may be other genetic loci participate in occurrence of both cancers. Therefore, it is rational and credible to discover the shared genetic background between ESCC and GC. Although we did not find connections between EAC and GC, it may be explained by population heterogeneity as EAC associated loci were reported in European ancestry. Moreover, only a few EAC risk loci were reported and used in the current study, which might be less representative for genetic risk of EAC based on WGS of EAC.

There are some limitations in our study. First, although the association between ESCC WGS and GC risk was significant, there were obvious heterogeneities among three studies for three SNPs. The associations became weak as considering the heterogeneity and should be treated in

caution. Second, we did not conduct subgroup analysis based on cardiac and non-cardiac GC, which limited our analysis on the different impact of these genetic factors among tumor subtype. Third, there are considerable difference between ESCC and GC, including environment and patient characteristics, which may have introduced potential bias on our results. Therefore, our results were preliminary and should be further evaluated in future studies.

In summary, we evaluated the genetic association between EC and GC, and found shared genetic susceptibility between ESCC and GC. In the future, more studies with larger sample sizes and multiple populations are needed to help detect the relationship between the genetic basis of EC and GC.

**Acknowledgements** We thank all the participants of the Nanjing/Beijing and the National Cancer Institute gastric cancer studies. This study was supported by grants from the National key research and development program of China (grant no. 2016YFC1302703); and National Natural Science Foundation of China (81521004, 81422042, 81872702).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Human rights statement and informed consent** All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and later versions. Informed consent to be included in the study, or the equivalent, was obtained from all patients.

## References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–386.
2. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. *Lancet*. 2016;388:2654–64.
3. Domper Arnal MJ, Ferrandez Arenas A, Lanas Arbeloa A. Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol*. 2015;21:7933–43.
4. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. *Lancet*. 2013;381:400–12.
5. Vaughan TL, Davis S, Kristal A, Thomas DB. Obesity, alcohol, and tobacco as risk factors for cancers of the esophagus and gastric cardia: adenocarcinoma versus squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev*. 1995;4:85–92.
6. Morita M, Kumashiro R, Kubo N, Nakashima Y, Yoshida R, Yoshinaga K, et al. Alcohol drinking, cigarette smoking, and the development of squamous cell carcinoma of the esophagus: epidemiology, clinical findings, and prevention. *Int J Clin Oncol*. 2010;15:126–34.
7. Brown LM, Silverman DT, Pottner LM, Schoenberg JB, Greenberg RS, Swanson GM, et al. Adenocarcinoma of the esophagus and esophagogastric junction in white men in the United States:



- alcohol, tobacco, and socioeconomic factors. *Cancer Causes Control*. 1994;5:333–40.
8. Lee CH, Wu DC, Lee JM, Wu IC, Goan YG, Kao EL, et al. Carcinogenetic impact of alcohol intake on squamous cell carcinoma risk of the oesophagus in relation to tobacco smoking. *Eur J Cancer*. 2007;43:1188–99.
  9. Runge TM, Abrams JA, Shaheen NJ. Epidemiology of Barrett's esophagus and esophageal adenocarcinoma. *Gastroenterol Clin North Am*. 2015;44:203–31.
  10. Abnet CC, Freedman ND, Hu N, Wang Z, Yu K, Shu XO, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat Genet*. 2010;42:764–7.
  11. Gharakhani P, Fitzgerald RC, Vaughan TL, Palles C, Gockel I, Tomlinson I, et al. Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: a large-scale meta-analysis. *Lancet Oncol*. 2016;17:1363–73.
  12. Shi Y, Hu Z, Wu C, Dai J, Li H, Dong J, et al. A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13.1. *Nat Genet*. 2011;43:1215–8.
  13. Cui R, Kamatani Y, Takahashi A, Usami M, Hosono N, Kawaguchi T, et al. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology*. 2009;137:1768–75.
  14. Wang LD, Zhou FY, Li XM, Sun LD, Song X, Jin Y, 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.
  15. Wu C, Hu Z, He Z, Jia W, Wang F, Zhou Y, 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.
  16. Abnet CC, Wang Z, Song X, Hu N, Zhou FY, Freedman ND, et al. Genotypic variants at 2q33 and risk of esophageal squamous cell carcinoma in China: a meta-analysis of genome-wide association studies. *Hum Mol Genet*. 2012;21:2132–41.
  17. Wu C, Kraft P, Zhai K, Chang J, Wang Z, Li Y, 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.
  18. Wu C, Wang Z, Song X, Feng XS, Abnet CC, He J, et al. Joint analysis of three genome-wide association studies of esophageal squamous cell carcinoma in Chinese populations. *Nat Genet*. 2014;46:1001–6.
  19. McKay JD, Truong T, Gaborieau V, Chabrier A, Chuang SC, Byrnes G, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet*. 2011;7:e1001333.
  20. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2011;9:179–81.
  21. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5:e1000529.
  22. Wing MR, Bourdon DM, Harden TK. PLC-epsilon: a shared effector protein in Ras-, Rho-, and G alpha beta gamma-mediated signaling. *Mol Interv*. 2003;3:273–80.
  23. Harden TK, Sondek J. Regulation of phospholipase C isozymes by ras superfamily GTPases. *Annu Rev Pharmacol Toxicol*. 2006;46:355–79.
  24. Hu H, Yang J, Sun Y, Yang Y, Qian J, Jin L, 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.
  25. Zhang H, Jin G, Li H, Ren C, Ding Y, Zhang Q, et al. Genetic variants at 1q22 and 10q23 reproducibly associated with gastric cancer susceptibility in a Chinese population. *Carcinogenesis*. 2011;32:848–52.
  26. Palmer AJ, Lochhead P, Hold GL, Rabkin CS, Chow WH, Lisowski J, et al. Genetic variation in C20orf54, PLCE1 and MUC1 and the risk of upper gastrointestinal cancers in Caucasian populations. *Eur J Cancer Prev*. 2012;21:541–4.
  27. Haddad SA, Ruiz-Narvaez EA, Haiman CA, Sucheston-Campbell LE, Bensen JT, Zhu Q, et al. An exome-wide analysis of low frequency and rare variants in relation to risk of breast cancer in African American Women: the AMBER Consortium. *Carcinogenesis*. 2016;37:870–7.
  28. Natrajan R, Mackay A, Lambros MB, Weigelt B, Wilkerson PM, Manie E, et al. A whole-genome massively parallel sequencing analysis of BRCA1 mutant oestrogen receptor-negative and -positive breast cancers. *J Pathol*. 2012;227:29–41.
  29. Pullamsetti SS, Banat GA, Schmall A, Szibor M, Pomagruk D, Hanze J, et al. Phosphodiesterase-4 promotes proliferation and angiogenesis of lung cancer by crosstalk with HIF. *Oncogene*. 2013;32:1121–34.
  30. Rahrmann EP, Collier LS, Knutson TP, Doyal ME, Kuslak SL, Green LE, et al. Identification of PDE4D as a proliferation promoting factor in prostate cancer using a Sleeping Beauty transposon-based somatic mutagenesis screen. *Cancer Res*. 2009;69:4388–97.
  31. Powers GL, Hammer KD, Domenech M, Frantskevich K, Malinowski RL, Bushman W, et al. Phosphodiesterase 4D inhibitors limit prostate cancer growth potential. *Mol Cancer Res*. 2015;13:149–60.
  32. Bottcher R, Dulla K, van Strijp D, Dits N, Verhoef EI, Baillie GS, et al. Human PDE4D isoform composition is deregulated in primary prostate cancer and indicative for disease progression and development of distant metastases. *Oncotarget*. 2016;7:70669–84.
  33. Henderson DJ, Byrne A, Dulla K, Jenster G, Hoffmann R, Baillie GS, et al. The cAMP phosphodiesterase-4D7 (PDE4D7) is downregulated in androgen-independent prostate cancer cells and mediates proliferation by compartmentalising cAMP at the plasma membrane of VCaP prostate cancer cells. *Br J Cancer*. 2014;110:1278–87.
  34. Chen, ZH, Xian, JF & Luo, LP Analysis of ADH1B Arg47His, ALDH2 Glu487Lys, and CYP4502E1 polymorphisms in gastric cancer risk and interaction with environmental factors. *Genet Mol Res*. 15 (2016).
  35. Duell EJ, Sala N, Travier N, Munoz X, Boutron-Ruault MC, Clavel-Chapelon F, et al. Genetic variation in alcohol dehydrogenase (ADH1A, ADH1B, ADH1C, ADH7) and aldehyde dehydrogenase (ALDH2), alcohol consumption and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. *Carcinogenesis*. 2012;33:361–7.
  36. Tsai ST, Wong TY, Ou CY, Fang SY, Chen KC, Hsiao JR, et al. The interplay between alcohol consumption, oral hygiene, ALDH2 and ADH1B in the risk of head and neck cancer. *Int J Cancer*. 2014;135:2424–36.
  37. Guo XF, Wang J, Yu SJ, Song J, Ji MY, Zhang JX, et al. Meta-analysis of the ADH1B and ALDH2 polymorphisms and the risk of colorectal cancer in East Asians. *Intern Med*. 2013;52:2693–9.
  38. Liu, P, Zhao, HR, Li, F, Zhang, L, Zhang, H, Wang, WR et al. Correlations of ALDH2 rs671 and C12orf36 rs4767364 polymorphisms with increased risk and prognosis of esophageal squamous cell carcinoma in the Kazak and Han populations in Xinjiang province. *J Clin Lab Anal*. 32 (2018).
  39. Gao Y, He Y, Xu J, Xu L, Du J, Zhu C, et al. Genetic variants at 4q21, 4q23 and 12q24 are associated with esophageal squamous cell carcinoma risk in a Chinese population. *Hum Genet*. 2013;132:649–56.
  40. Yu, C, Guo, Y, Bian, Z, Yang, L, Millwood, IY, Walters, RG et al. Association of low-activity ALDH2 and alcohol

- consumption with risk of esophageal cancer in Chinese adults: a population-based cohort study. *Int J Cancer*. (2018).
41. Hidaka A, Sasazuki S, Matsuo K, Ito H, Sawada N, Shimazu T, et al. Genetic polymorphisms of ADH1B, ADH1C and ALDH2, alcohol consumption, and the risk of gastric cancer: the Japan Public Health Center-based prospective study. *Carcinogenesis*. 2015;36:223–31.
  42. Yang S, Lee J, Choi IJ, Kim YW, Ryu KW, Sung J, et al. Effects of alcohol consumption, ALDH2rs671 polymorphism, and *Helicobacter pylori* infection on the gastric cancer risk in a Korean population. *Oncotarget*. 2017;8:6630–41.
  43. Zhang LQ, Song X, Zhao XK, Huang J, Zhang P, Wang LW, et al. Association of genotypes of rs671 within ALDH2 with risk for gastric cardia adenocarcinoma in the Chinese Han population in high- and low-incidence areas. *Cancer Biol Med*. 2017;14:60–65.
  44. Shin CM, Kim N, Cho SI, Kim JS, Jung HC, Song IS. Association between alcohol intake and risk for gastric cancer with regard to ALDH2 genotype in the Korean population. *Int J Epidemiol*. 2011;40:1047–55.
  45. Testino G, Leone S, Patussi V, Scafato E. [Alcohol, cardiovascular prevention and cancer]. *Recent Prog Med*. 2014;105:144–6.
  46. Haas SL, Ye W, Lohr JM. Alcohol consumption and digestive tract cancer. *Curr Opin Clin Nutr Metab Care*. 2012;15:457–67.