




Research Paper

SMADs binding site polymorphisms rs9911630 is associated with susceptibility but not prognosis of gastric cancer: a case control study

Liyang Liu^{1*}, Ming Lu^{1*}, Xi Gu^{1*}, Xiang Ma¹, Jiayi Feng¹, Yang Cao¹, Weida Gong², Qinghong Zhao¹, Fulin Qiang³

1. Department of General Surgery, the Second Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China.
2. Department of General Surgery, Yixing Tumor Hospital, Yixing, Jiangsu, China.
3. Department of Core Laboratory, Nantong Tumor Hospital, Nantong, Jiangsu, China.

*Equal contributors and co-first authors.

 Corresponding authors: Qinghong Zhao. Department of General Surgery, The Second Affiliated Hospital of Nanjing Medical University, 121 Jiangjia Yuan Road, Nanjing 210011, Jiangsu, China. E-mail: zhaoqinghong@njmu.edu.cn; Fulin Qiang. Department of Core Laboratory, Nantong Tumor Hospital, Nantong, China. E-mail: qiangfulin@126.com.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2019.09.08; Accepted: 2020.05.02; Published: 2020.06.01

Abstract

Background: Single nucleotide polymorphisms (SNPs) in transcription factor binding sites (TFBS) can change their binding strength, affecting the function of transcription factors (TFs). Small mother against decapentaplegic (SMAD) proteins are known as a family of TFs involved in tumorigenesis. We performed this study to investigate whether SNPs in SMADs binding sites affect the susceptibility or prognosis of gastric cancer (GC).

Methods: Using bioinformatics tools, we focused on the association between rs9911630 polymorphism and GC. We performed this case-control study in 1275 GC patients and 1426 cancer-free subjects using TaqMan allelic discrimination method.

Results: We found that rs9911630 A>G polymorphism was associate to an increased risk of gastric cancer (adjusted OR for additive model = 1.16; 95% CI = 1.03-1.30). Furthermore, we assess whether rs9911630 polymorphism affected the prognosis of GC. However, no significant association was discovered between rs9911630 A>G polymorphism and overall survival time of GC patients (HR for additive model = 1.01; 95%CI = 0.88-1.15).

Conclusions: Our results suggested that rs9911630 polymorphism in SMADs target site might influence susceptibility but not prognosis of gastric cancer.

Key words: rs9911630, polymorphisms, susceptibility, gastric cancer

Introduction

Gastric cancer (GC) is the fourth most common cancer and second dominant cause of cancer-related death worldwide [1,2]. In spite of major improvements in diagnosis and treatment, the 5-year survival rate of GC is still less than 25% [3]. Therefore, it is urgently required to identify a new way for predicting GC susceptibility and prognosis [4-6]. Both environmental and genetic factors are involved in etiology of GC. Environmental risk factors such as older age, *Helicobacter pylori* infection and tobacco

smoking nowadays are well-known for their role in GC [7]. Pathogenetic mechanisms in GC are still being debated but in recent years, a number of single nucleotide polymorphisms (SNPs) have been found playing a vital role in gastric carcinogenesis [8].

Transcription factor (TF) dysregulation, playing a vital role in abnormal gene expression, is a hallmark of many cancers [9,10]. The genomic locations of TF binding at specific locus have functional consequences with respect to the binding ability of TF

[11]. SNPs seating in transcriptional factors binding sites (TFBS) may conclusively influence the binding ability and modulate individual cancer susceptibility [12-14]. In addition, studies indicated that SNPs may modify the methylation level of gene promoter regions, interfering with TF binding, which in turn leads to abnormalities of gene transcription [15,16]. The identification of these SNPs that represent a functional link with methylation sites provided functional insight into the potential mechanism by which genetic variants involved in etiology of tumor.

Small mother against decapentaplegic (SMAD) proteins, as a family of transcription factors, are expressed broadly in the body tissues [17]. SMAD proteins act as mediators of transforming growth factor-beta (TGF- β) signaling, which is one of the most important tumor suppressor pathways [18]. SMADs translocate signals from the cell surface to the nucleus, regulating TGF- β superfamily-dependent gene expression [19]. The TGF- β /SMAD signaling pathway was found to regulate cell growth and promotes apoptosis of epithelial cells, and participate in angiogenesis [20]. Accumulating evidence indicated that components of this pathway are involved in a large range of cancers [21-23]. Function of this signaling pathway may be influenced when a genetic variant occurs in the SMADs' binding site. We evaluated the associations between these SNPs and GC susceptibility using GWAS data. Among all these eligible SNPs, we found rs9911630 A>G could affect the methylation level of CpG sites in promoter regions of three genes. So, we selected rs9911630 and evaluated its effect on the susceptibility and prognosis of GC in this study.

Methods

Study population

There were 1,275 GC cases and 1,426 age- and sex-matched cancer-free controls covered in our

study. All cases were supported by the Cancer Clinical Research Base of Nanjing Medical University between March 2006 and May 2013. Only histologically confirmed GC patients were included. Exclusion criteria included secondary GC or metastasized cancer from other organs. In addition, patients that received neoadjuvant chemotherapy or radiotherapy before surgery were excluded. All control subjects were randomly enrolled at the same period when they sought physical examinations in the same hospital. The controls were frequency-match to cases on age (± 5 years) and sex. All patients enrolled in this study were genetically unrelated ethnic Han Chinese. The study was authorized by the institutional review board of Nanjing Medical University. Every participant enrolled in this study signed an informed consent.

SNPs selection

SNPs located in SMADs binding sites were searched according to genotype data of genome-wide association studies (GWASs). Then we evaluated the associations between these SNPs and GC susceptibility using GWAS data and identified eligible SNPs with a standard of $P < 0.01$ and minor allele frequency (MAF) > 0.05 . A total of 556 relevant SNPs were obtained from GWAS datasets, and after the process of our selection, 8 eligible SNPs were taken into further consideration (**Table 1**). We would like to focus on SNPs acting as methylation Quantitative Trait Loci (meQTLs) in the surrounding region. The level 3 Human Methylation 450 and Level 2 SNP Array data of gastric adenocarcinoma were downloaded from the The Cancer Genome Atlas (TCGA) database. We tested the methylation status of CpG sites situated within 10000 bases range of each SNP by meQTL analysis. Finally, rs9911630 were enrolled in further study.

Table 1. Characteristics of the selected SNPs

SNP	Gene	Allele	MAF	OR (95% CI) ^a	meQTL (risk allele association, P)
rs17707882	MYO10	C>T	0.148	0.84 (0.73-0.96)	cg18061395, $P = 1.31 \times 10^{-4}$ (decreased) cg24556395, $P = 4.22 \times 10^{-5}$ (decreased)
rs9353563	CNR1	A>G	0.246	1.11 (1.01-1.24)	cg02436141, $P = 2.63 \times 10^{-4}$ (decreased) cg18703951, $P = 4.34 \times 10^{-3}$ (decreased) cg08458400, $P = 1.85 \times 10^{-2}$ (decreased) cg23276695, $P = 3.57 \times 10^{-2}$ (increased)
rs1569836	AGPAT4	A>G	0.275	1.14 (1.03-1.26)	cg02031769, $P = 2.25 \times 10^{-3}$ (decreased) cg11697870, $P = 1.73 \times 10^{-2}$ (decreased)
rs10514486	SLC36A4	T>C	0.397	0.90 (0.82-0.99)	cg15554438, $P = 2.96 \times 10^{-2}$ (decreased)
rs9911630	LINC0091 (lncRNA), NBR1, BRCA1	A>G	0.334	1.11 (1.01-1.23)	cg10047753, $P = 2.68 \times 10^{-30}$ (LINC00910, decreased) cg23758822, $P = 2.43 \times 10^{-26}$ (LINC00910, decreased) cg05368731, $P = 1.70 \times 10^{-16}$ (NBR1, increased)

SNP	Gene	Allele	MAF	OR (95% CI) ^a	meQTL (risk allele association, P)
					cg19454999, $P = 1.68 \times 10^{-12}$ (NBR1, increased)
					cg25072359, $P = 8.45 \times 10^{-7}$ (NBR1, increased)
					cg25918947, $P = 1.19 \times 10^{-5}$ (NBR1, increased)
					cg01879757, $P = 8.50 \times 10^{-5}$ (BRCA1, increased)
					cg25067162, $P = 1.42 \times 10^{-5}$ (BRCA1, increased)
rs618443	CCNY	G>A	0.375	1.16 (1.06-1.28)	cg05845615, $P = 2.65 \times 10^{-2}$ (decreased)
rs4940826	LMAN1	A>G	0.21	1.12 (1.00-1.26)	cg25361621, $P = 1.55 \times 10^{-4}$ (increased)
rs10423232	ANKLE1	T>C	0.324	0.89 (0.81-0.99)	cg00433770, $P = 2.39 \times 10^{-3}$ (increased)

^aAdjusted by age and sex in logistic additive analysis.

SNPs genotyping

We isolated genomic DNA from peripheral blood. The selected SNPs were genotyped using TaqMan allelic discrimination assay on the ABI 7900HT Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). For confirmation, 10% of the samples were selected to be genotyped again, and the results were in consistent with the first assay. The structure of primers and probes are as follows: forward primer: 5'- GCTCTCTAAGGTCCCTTCTCA TTG-3', reverse primer: 5' -GCACAAGTGACCGATG GGTA-3', and probes: FAM: AAGCACAGTGCAT GGA, HEX-AAGCACAGCGCATG.

Statistical analysis

We assessed the differences in demographic factors by Student's t test and Pearson's chi-squared (χ^2) test. Hardy-Weinberg equilibrium (HWE) of the controls was assessed by a goodness-of-fit χ^2 test. The ORs and 95% CIs were calculated to estimate associations between these SNPs and GC susceptibility. Variables of age and sex were used as covariates adjusted for the association analysis. We used multiple inheritance models to estimate the significance of SNP rs9911630. Kaplan-Meier method and log-rank test were applied to evaluate the associations between survival time and the SNP rs9911630. Mean survival time was provided when the median survival time (MST) could not be calculated. We performed Univariate or multivariate Cox regression analysis to calculate crude or adjusted hazard ratios (HRs) and 95% CIs. $P < 0.05$ for two-side Student's t test was considered statistically significant when analyzing the promoter activity. All statistical analyses were carried out using SAS software (version 9.1.3; SAS Institute Inc., Cary, NC, USA).

Results

Association between SNP rs9911630 and GC risk from publicly databases

We downloaded publicly available GWAS datasets from dbGaP database. We used additive

model to evaluate the association between SNP rs9911630 and GC risk. As a result, we found rs9911630 A>G polymorphism were significantly associated with GC risk (adjusted OR = 1.11, 95% CI = 1.01-1.23, $P < 0.01$). We performed meQTL analysis based on TCGA datasets to test whether these SNPs are associated with the methylation status of CpG sites situated nearby. As shown in **Table 1**, rs9911630 A>G was related to methylation level of CpG sites in promoter regions of three genes (the neighbor of *brca1* gene, the breast and ovarian cancer susceptibility gene 1 and long intergenic non-coding RNA 910). rs9911630 G allele was related to the decreased methylation level of cg10047753 and cg23758822 ($P = 2.68 \times 10^{-30}$ for cg10047753 and $P = 2.43 \times 10^{-26}$ for cg23758822). Besides, rs9911630 G allele was related to the increased methylation level of other six CpG sites ($P = 1.70 \times 10^{-16}$ for cg05368731, $P = 1.68 \times 10^{-12}$ for cg19454999, $P = 8.45 \times 10^{-7}$ for cg25072359, $P = 1.19 \times 10^{-5}$ for cg25918947, $P = 8.50 \times 10^{-5}$ for cg01879757, $P = 1.42 \times 10^{-5}$ for cg25067162).

Characteristics of cases and controls

Then we perform a case-control study to further evaluate the associations between SNP rs9911630 and GC susceptibility using our samples. In this study, no remarkable difference was found among cases and controls in the distributions of age ($P = 0.324$) and sex ($P = 0.358$). Clinicopathological characteristics of case-control studies were summarized in **Table 2**. Of these cases, there were 61.3% non-cardia gastric cancer patients, and 33.6% cardia gastric cancer patients; 682 (61.4%) had lymph node metastasis and 167 (15.1%) existed distant metastasis. In addition, all the cases were identified to the TNM stage in accordance with the 6th edition staging manual of the American Joint Committee on Cancer (AJCC). TNM stage I, II, III, and IV were with the percentage of 23.1%, 24.6%, 35.5%, and 16.8%, respectively.

Association of rs9911630 polymorphism with gastric cancer risk

Genotype distributions rs9911630 among the patients and controls were shown in **Table 3**. The genotype frequencies were agreed with the Hardy-

Weinberg equilibrium ($P = 0.1166$). Different inheritance models were used and the results indicated that rs9911630 A>G polymorphism were significantly associated with GC risk in additive models (adjusted OR = 1.16, 95% CI = 1.03-1.30, $P = 0.012$); codominant model (adjusted OR for GG genotype = 1.39, 95% CI = 1.09-1.78, $P = 0.009$) and recessive model (adjusted OR = 1.32, 95% CI = 1.05-1.66, $P = 0.020$). As result, rs9911630 G allele was a potential risk allele for GC. The main findings of this case-control study were consistent with analysis based on publicly databases.

Table 2. Characteristics of study subjects

Variables	Cases		Controls		P^a
	N=1275	%	N=1426	%	
Age (years) (mean±SD)	63.1±10.7		63.3±11.0		0.595
Sex					
Male	880	69.0	963	67.5	0.347
Female	392	30.8	463	32.5	
NA	3	0.2			
Tumor site					
Cardia	403	33.6			
Non-cardia	734	61.3			
Both	61	5.1			
NA	77				
Histological types					
Intestinal	513	45.6			
Diffuse	612	54.4			
NA	150				
Depth of invasion					
Tis	1	0.1			
T1	170	15.2			
T2	169	15.1			
T3	575	51.5			
T4	204	18.1			
NA	158				
Lymph node metastasis					
N0	428	35.6			
N1/N2/N3	682	61.4			
NA	165				
Distant metastasis					
M0	941	84.9			
M1	167	15.1			
NA	167				
TNM					
I	267	23.1			
II	284	24.6			
III	410	35.5			
IV	194	16.8			
NA	120				

Two-sided student t test for the frequency distributions of age between the cases and controls. Two-sided χ^2 test for the frequency distributions of sex between the cases and controls.

Table 3. Association of rs9911630 polymorphism with gastric cancer risk

	Genotype	Cases/controls	OR (95% CI)	Adjusted OR (95% CI) ^a	P^b
Additive model	AA	497/603	1.15 (1.03-1.29)	1.16 (1.03-1.30)	0.012
	AG	603/669			
	GG	175/154			
Codominant model	AA	497/603	1.00	1.00	0.227
	AG	603/669	1.09 (0.93-1.29)	1.11 (0.94-1.30)	

Stratified analysis of SNP rs9911630 and GC risk

We analyzed the effects of rs9911630 polymorphism on GC risk stratified in accordance with different clinical variables. As shown in **Table 4**, we found association between rs9911630 G allele and increased risk of GC among subgroup of non-cardia (adjusted OR = 1.21, 95% CI = 1.06-1.38, $P = 0.004$) and histological types of diffuse (adjusted OR = 1.27, 95% CI = 1.11-1.47, $P = 0.001$). Significant risk effect was not observed in subgroups of different depth of invasion, lymph node metastasis, distant metastasis or TNM stages.

SNP rs9911630 polymorphism and gastric cancer survival

Since rs9911630 G allele was a potential risk allele for GC, we would like to assess the prognostic value of SNP rs9911630 on GC patients. This study comprised 933 patients with gastric cancer and overall survival was the end point. Characteristics and clinical features of subjects were summarized in **Table 5**. Histology, the depth of invasion, lymph node status, distant metastasis and TNM stage were factors affecting the survival time of GC patients (log-rank $P < 0.05$).

We used log-rank test to evaluate the effect of rs9911630 A>G on overall survival time in GC patients. However, as the presented in **Table 6**, significant association was not observed between the rs9911630 polymorphism and overall survival time in additive model (log-rank $P = 0.691$), dominant model (log-rank $P = 0.630$) or recessive model (log-rank $P = 0.612$). To further assess the association between the rs9911630 and survival of patients with GC, we performed subgroup analyses by clinical characteristics under dominant model. There was no prominent association between SNP rs9911630 polymorphism and survival time when stratified by age, sex, tumor site, histology, depth of invasion, lymph node metastasis, distant metastasis, TNM stage and chemotherapy (**Table 7**). As a result, we did not find significant association between rs9911630 and GC prognosis in the present study.

	Genotype	Cases/controls	OR (95% CI)	Adjusted OR (95% CI) ^a	P ^a
Dominant model	GG	175/154	1.38 (1.08-1.77)	1.39 (1.09-1.78)	0.009
	AA	497/603	1.00	1.00	
	AG/GG	778/823	1.15 (0.98-1.34)	1.16 (0.99-1.35)	0.062
Recessive model	AA/AG	1100/1272	1.00	1.00	
	GG	175/154	1.32 (1.04-1.66)	1.32 (1.05-1.66)	0.020

^aAdjusted by age and sex in logistic additive analysis.

Table 4. Associations between rs9911630 genotypes and clinical characteristics of GC

Variables	OR (95% CI)	Adjusted OR (95% CI) ^a	P ^a
Controls	1.00	1.00	
Tumor site			
Cardia	1.01 (0.85-1.19)	1.02 (0.86-1.20)	0.835
Non-cardia	1.21 (1.06-1.38)	1.21 (1.06-1.38)	0.004
Histological types			
Diffuse	1.26 (1.10-1.45)	1.27 (1.11-1.47)	0.001
Intestinal	1.02 (0.88-1.18)	1.02 (0.87-1.19)	0.825
Depth of invasion			
T1	1.25 (0.99-1.58)	1.25 (0.99-1.59)	0.063
T2	1.10 (0.87-1.40)	1.12 (0.88-1.42)	0.367
T3	1.08 (0.94-1.25)	1.08 (0.94-1.25)	0.283
T4	1.12 (0.90-1.40)	1.13 (0.90-1.41)	0.291
Lymph node metastasis			
N0	1.10 (0.94-1.30)	1.11 (0.95-1.31)	0.196
N1/N2/N3	1.12 (0.98-1.29)	1.12 (0.98-1.29)	0.097
Distant metastasis			
M0	1.10 (0.97-1.24)	1.10 (0.97-1.24)	0.128
M1	1.22 (0.97-1.53)	1.25 (0.99-1.58)	0.056
TNM stages			
I+II	1.12 (0.97-1.30)	1.12 (0.97-1.30)	0.123
III+IV	1.12 (0.97-1.29)	1.13 (0.98-1.30)	0.098

^aAdjusted by age and sex in logistic additive analysis.

Discussion

As a family of transcription factors, SMADs may impact regulation of target genes, participating in cancer-related biological processes [24,25]. The abnormal expression of SMADs was found in several human malignancies, including GC [26,27]. Wu et al. [28] have reported that genetic variations in SMAD4 gene are related to GC susceptibility. Recently, many evidences have been documented between TFBS and GC pathogenesis. Hence, SNPs in SMADs binding sites are expected to become risk markers for GC.

In this study, we studied one SNP (rs9911630 A>G) lying in the binding site of SMADs to explore its association with GC susceptibility and survival. Firstly, we found SNP rs9911630 was related to GC risk according to publicly databases. Then we perform a study of 1,275 cases and 1,426 controls to further evaluate the associations between SNP rs9911630 and GC susceptibility. Results showed that rs9911630 can influence the risk of GC, and individuals with the rs9911630 variant genotypes (GG) had observably increased GC risk compared with those with the AA/AG genotypes. However, significantly association was not observed between rs9911630 A>G

polymorphism and overall survival time of GC patients. In addition, we found rs9911630 G allele was associated with increased risk of non-cardia GC but not cardia GC. Many SNPs have been reported to be susceptibility locus specific for cardia GC or for non-cardia GC [29]. Gastric cardia carcinoma differs from non-cardia carcinoma in epidemiological characteristics, etiology and clinical features. Risk factors for gastric cardia adenocarcinomas also differ between these two main sub-locations of GC. Studies in western populations have put forward that cardia adenocarcinomas are more similar to esophageal adenocarcinomas [30]. SNPs in a locus on chromosome 10q23 in the *PLCE1* gene were reported to have strong association with gastric cardia adenocarcinoma and esophageal squamous cell carcinoma, but no association with gastric non-cardia adenocarcinoma [31]. These findings suggested that identification of phenotype-specific genetic susceptibility loci could improve understanding of different subtypes of GC, which in turn is important for early detection, diagnosis and treatment of this malignancy.

In the present study, we did not perform functional study to estimate the role of rs9911630 polymorphism in the current study. However, we performed meQTL analysis based on TCGA datasets, and as a result, we found that rs9911630 A>G was associated with methylation level of CpG sites in promoter regions of three genes (the neighbor of *brca1* gene, the breast and ovarian cancer susceptibility gene 1 and long intergenic non-coding RNA 910). DNA methylation plays an important role in modulating the transcription of mammalian genomes by blocking the binding of transcription factors. Previous studies have demonstrated that SNPs may modify the methylation level of CpG sites or influence the generation of new CpG sites, which changes the status of genes' methylation and regulate gene expression [32-34]. Accordingly, we speculated that the SNP rs9911630 could influence the binding ability of SMADs and change methylation level of the gene promoter regions nearby, which in turn leads to the influence of gene outputs. Despite the exact mechanism remained to be elucidated, these functions of SNP rs9911630 may play roles in gastric carcinogenesis.

Table 5. Patients' characteristics and clinical features

Variables	Patients (n=933)	Deaths (n=439)	MST (months)	Log-rank <i>p</i>	Adjusted HR (95% CI)
Age					
≤60	436	201	90.1	0.285	1.00
>60	497	235	60.0		1.11 (0.92-1.34)
Sex					
Male	718	332	75.5	0.412	1.00
Female	215	104	64.3		1.10 (0.88-1.37)
Site					
Cardia	356	165	66.9	0.580	1.00
Non-cardia	577	271	71.0		1.06 (0.87-1.28)
Histology					
Diffuse	536	280	51.3	0.001	1.00
Intestinal	397	156	57.6 ^a		0.72 (0.59-0.88)
Depth of invasion					
T1	149	45	48.7 ^a	<0.001	1.00
T2	199	83	90.1		1.54 (1.07-2.21)
T3	540	284	49.2		2.15 (1.57-2.95)
T4	45	27	26.9		2.77 (1.72-4.46)
Lymph node metastasis					
N0	372	130	83.1 ^a	<0.001	1.00
N1-N3	561	306	44.4		1.87 (1.52-2.29)
Distant metastasis					
M0	875	401	75.5	0.003	1.00
M1	58	35	27.5		1.67 (1.18-2.36)
TNM stage					
I+II	259	87	60.8 ^a	<0.001	1.00
III+IV	578	294	54.5		1.79 (1.41-2.27)
Chemotherapy					
No	629	299	75.1	0.728	1.00
Yes	304	137	61.5		1.04 (0.85-1.27)

^aMean survival time was provided when MST could not be calculated.

Table 6. Association between rs9911630 and overall survival of GC

SNP	Genetic models	Genotypes	All cases	Deaths	MST	Log-rank <i>p</i>	HR (95% CI) ^a
rs9911630	addictive	AA	350	163	82.1	0.691	1.01 (0.88-1.15)
		AG	439	210	60.0		
		GG	144	63	66.9		
	dominant	AA	350	163	82.1	0.630	1.00
		AG/GG	583	273	63.5		
	recessive	AA/AG	789	373	71.0	0.612	1.00
	GG	144	63	66.9		0.93 (0.71-1.21)	

^aAdjusted for age and sex.

Trying the best of ourselves with existing materials, this is the first study exploring the association between the SNP rs9911630 polymorphism of SMADs binding site and susceptibility and prognosis of GC. However, the current study was subject to limitations. Firstly, some environmental factors like smoking, drinking and *Helicobacter pylori* infection play vital roles in gastric carcinogenesis, but due to the unavailability of detailed information in some of the study subjects, we did not perform a further analysis to investigate the gene-environment interaction. Secondly, sample size of the current study is small, making analysis less reliable than if a large sample had been available. Thirdly, functional study was not operated to estimate the role of rs9911630 polymorphism.

Table 7. Subgroup analyses of association between rs9911630 polymorphisms and GC survival

Variables	Genotype (deaths/patients)		HR (95% CI) ^a
	AA	AG/GG	
Total	163/350	273/583	1.06(0.87-1.29)
Age			
≤60	154/235	124/201	0.91(0.69-1.21)
>60	156/262	149/235	1.19(0.92-1.56)
Sex			
Male	239/386	203/332	1.02(0.82-1.28)
Female	71/111	70/104	1.183(0.785-1.783)
Site			
Cardia	121/191	998/165	0.953(0.698-1.301)
Non-Cardia	189/306	175/271	1.124(0.876-1.443)
Histology			
Diffuse	159/256	166/280	0.957(0.753-1.217)
Intestinal	151/241	107/156	1.333(0.950-1.871)
Depth of Invasion			
T1	67/104	29/45	0.974(0.524-1.809)

Variables	Genotype (deaths/patients)		HR (95% CI) ^a
	AA	AG/GG	
T2	68/117	56/82	1.422(0.891-2.267)
T3	164/258	172/282	0.975(0.767-1.240)
T4	11/18	16/27	1.140(0.508-2.558)
Lymph node metastasis			
N0	143/242	81/130	1.157(0.811-1.651)
N1-N3	167/255	192/306	0.986(0.781-1.246)
Distant metastasis			
M0	293/474	248/401	1.053(0.860-1.289)
M1	17/23	25/35	1.258(0.571-2.771)
TNM stage			
I-II	104/172	56/87	1.187(0.762-1.850)
III-IV	177/284	181/294	1.032(0.815-1.307)
Chemotherapy			
No	201/330	181/299	1.045(0.828-1.318)
Yes	109/167	92/137	1.104(0.771-1.579)

^aAdjusted for age and sex.

Conclusions

In conclusion, our results suggested that rs9911630 polymorphism in SMADs target site might influence susceptibility but not prognosis of GC in the Chinese populations. Meanwhile, methylation level of the nearby gene promoter regions could be changed according to the polymorphism rs9911630, which might influence the expression of these genes. Larger, well-designed epidemiologic and functional studies are still needed to prove these findings.

Abbreviations

SNPs: single nucleotide polymorphisms (SNPs); TFBS: transcription factor binding sites; TFs: transcription factors; SMAD: small mother against decapentaplegic; GC: gastric cancer; TGF- β : transforming growth factor-beta; GWASs: data of genome-wide association studies; MAF: minor allele frequency; meQTLs: methylation Quantitative Trait Loci; TCGA: The Cancer Genome Atlas; HWE: Hardy-Weinberg equilibrium; MST: median survival time.

Acknowledgements

Declarations

Ethics approval and consent to participate: the study was approved by the institutional review board of Nanjing Medical University, and an informed consent was signed by all participants.

Consent for publication: not applicable.

Availability of data and material: the datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Funding: the design of this study was supported by grants from the National Natural Science Foundation of Jiang Su Province of China (SBK2017022379), the collection of study population was supported by General Project of Nanjing Medical

Technology Development Fund (YKK16224), Science and Technology Development Fund of Nanjing Medical University (2014NJMU140), Youth talent support program of Nanjing City during the 13th Five-Year Plan Period (QRX17208), the interpretation of data and writing of the manuscript was supported by General Project of Nanjing Medical Technology Development Fund (YKK17208) Science and Technology Development Fund of Nanjing Medical University (2016NJMU040).

Authors' Contributions

Liyang Liu and Xi Gu: writing - original draft, Jiayi Feng: data curation, Weida Gong: formal analysis, Ming Lu and Xiang Ma: project administration. Qinghong Zhao and Fulin Qiang: writing - review & editing.

Acknowledgements: we thank members of our laboratory for helpful comments and discussion.

Competing Interests

The authors have declared that no competing interest exists.

References

- Ferro A, Peleteiro B, Malvezzi M, Bosetti C, Bertuccio P, Levi F, et al. Worldwide trends in gastric cancer mortality (1980-2011), with predictions to 2015, and incidence by subtype. *European journal of cancer*. 2014; 50: 1330-44.
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: a cancer journal for clinicians*. 2011; 61: 69-90.
- Camargo MC, Kim WH, Chiaravalli AM, Kim KM, Corvalan AH, Matsuo K, et al. Improved survival of gastric cancer with tumour Epstein-Barr virus positivity: an international pooled analysis. *Gut*. 2014; 63: 236-43.
- Dicken BJ, Bigam DL, Cass C, Mackey JR, Joy AA, Hamilton SM. Gastric adenocarcinoma: review and considerations for future directions. *Ann Surg*. 2005; 241: 27-39.
- Khan FA, Shukla AN. Pathogenesis and treatment of gastric carcinoma: "an up-date with brief review". *J Cancer Res Ther*. 2006; 2: 196-9.
- Catalano V, Labianca R, Beretta GD, Gatta G, de Braud F, Van Cutsem E. Gastric cancer. *Crit Rev Oncol Hematol*. 2009; 71: 127-64.
- Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol Biomarkers Prev*. 2014; 23: 700-13.
- Saeki N, Ono H, Sakamoto H, Yoshida T. Genetic factors related to gastric cancer susceptibility identified using a genome-wide association study. *Cancer Sci*. 2013; 104: 1-8.
- Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell*. 2013; 152: 1237-51.
- Bhagwat AS, Vakoc CR. Targeting Transcription Factors in Cancer. *Trends in cancer*. 2015; 1: 53-65.
- Cusanovich DA, Pavlovic B, Pritchard JK, Gilad Y. The functional consequences of variation in transcription factor binding. *PLoS genetics*. 2014; 10: e1004226.
- Hildebrandt MA, Yang H, Hung MC, Izzo JG, Huang M, Lin J, et al. Genetic variations in the PI3K/PTEN/AKT/mTOR pathway are associated with clinical outcomes in esophageal cancer patients treated with chemoradiotherapy. *J Clin Oncol*. 2009; 27: 857-71.
- Wang S, Wu S, Zhu H, Ding B, Cai Y, Ni J, et al. PSCA rs2294008 polymorphism contributes to the decreased risk for cervical cancer in a Chinese population. *Sci Rep*. 2016; 6: 23465.
- Liu Y, Walavalkar NM, Dozmorov MG, Rich SS, Civelek M, Guertin MJ. Identification of breast cancer associated variants that modulate transcription factor binding. *PLoS genetics*. 2017; 13: e1006761.
- van Eijk KR, de Jong S, Strengman E, Buijzer-Voskamp JE, Kahn RS, Boks MP, et al. Identification of schizophrenia-associated loci by combining DNA methylation and gene expression data from whole blood. *Eur J Hum Genet*. 2015; 23: 1106-10.
- Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics*. 2010; 5: 578-82.

17. Massague J, Blain SW, Lo RS. TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell*. 2000; 103: 295-309.
18. Chandrasinghe P, Cereser B, Moorghen M, Al Bakir I, Tabassum N, Hart A, et al. Role of SMAD proteins in colitis-associated cancer: from known to the unknown. *Oncogene*. 2018; 37: 1-7.
19. Derynck R, Zhang Y, Feng XH. Smads: transcriptional activators of TGF-beta responses. *Cell*. 1998; 95: 737-40.
20. Gordon KJ, Blobel GC. Role of transforming growth factor-beta superfamily signaling pathways in human disease. *Biochimica et biophysica acta*. 2008; 1782: 197-228.
21. Slattery ML, Lundgreen A, Herrick JS, Wolff RK, Caan BJ. Genetic variation in the transforming growth factor-beta signaling pathway and survival after diagnosis with colon and rectal cancer. *Cancer*. 2011; 117: 4175-83.
22. Wang C, Li Y, Zhang H, Liu F, Cheng Z, Wang D, et al. Oncogenic PAK4 regulates Smad2/3 axis involving gastric tumorigenesis. *Oncogene*. 2014; 33: 3473-84.
23. Xiangming C, Natsugoe S, Takao S, Hokita S, Ishigami S, Tanabe G, et al. Preserved Smad4 expression in the transforming growth factor beta signaling pathway is a favorable prognostic factor in patients with advanced gastric cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2001; 7: 277-82.
24. Wrana JL. Regulation of Smad activity. *Cell*. 2000; 100: 189-92.
25. Morikawa M, Koinuma D, Miyazono K, Heldin CH. Genome-wide mechanisms of Smad binding. *Oncogene*. 2013; 32: 1609-15.
26. Miyaki M, Iijima T, Konishi M, Sakai K, Ishii A, Yasuno M, et al. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene*. 1999; 18: 3098-103.
27. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997; 390: 465-71.
28. Wu DM, Zhu HX, Zhao QH, Zhang ZZ, Wang SZ, Wang ML, et al. Genetic variations in the SMAD4 gene and gastric cancer susceptibility. *World journal of gastroenterology*. 2010; 16: 5635-41.
29. 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. *Nature genetics*. 2011; 43: 1215-8.
30. Hu N, Wang Z, Song X, Wei L, Kim BS, Freedman ND, et al. Genome-wide association study of gastric adenocarcinoma in Asia: a comparison of associations between cardia and non-cardia tumours. *Gut*. 2016; 65: 1611-8.
31. 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. *Nature genetics*. 2010; 42: 764-7.
32. Moser D, Ekawardhani S, Kumsta R, Palmason H, Bock C, Athanassiadou Z, et al. Functional analysis of a potassium-chloride co-transporter 3 (SLC12A6) promoter polymorphism leading to an additional DNA methylation site. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*. 2009; 34: 458-67.
33. Shield AJ, Murray TP, Cappello JY, Coggan M, Board PG. Polymorphisms in the human glutathione transferase Kappa (GSTK1) promoter alter gene expression. *Genomics*. 2010; 95: 299-305.
34. Raptis S, Mrkonjic M, Green RC, Pethe VV, Monga N, Chan YM, et al. MLH1 -93G>A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. *Journal of the National Cancer Institute*. 2007; 99: 463-74.