Associations of genotypes and haplotypes of \textit{IL-17} with risk of gastric cancer in an eastern Chinese population

Fei Zhou$^{1,2,3}$, Li-Xin Qiu$^{1,2}$, Lei Cheng$^{1,2}$, Meng-Yun Wang$^{1,2}$, Jin Li$^{2}$, Meng-Hong Sun$^{4}$, Ya-Jun Yang$^{5,6}$, Jiu-Cun Wang$^{5,6}$, Li Jin$^{5,6}$, Ya-Nong Wang$^{7}$, Qing-Yi Wei$^{1,8}$

$^1$Cancer Institute, Collaborative Innovation Center for Cancer Medicine, Fudan University Shanghai Cancer Center, Shanghai, China
$^2$Department of Medical Oncology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China
$^3$Department of Oncology, Shanghai Jiaotong University Affiliated Shanghai First People’s Hospital, Shanghai, China
$^4$Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China
$^5$Ministry of Education Key Laboratory of Contemporary Anthropology, State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China
$^6$Fudan-Taizhou Institute of Health Sciences, Taizhou, Jiangsu, China
$^7$Department of Gastric Cancer and Soft Tissue Sarcoma Surgery, Fudan University Shanghai Cancer Center, Shanghai, China
$^8$Duke Cancer Institute, Duke University Medical Center, Durham, NC, USA

Correspondence to: Ya-Nong Wang, email: wangyn1111@hotmail.com
Qing-Yi Wei, email: weiqingyi@yahoo.com

Keywords: interleukin-17, genetic variants, gastric cancer, susceptibility, molecular epidemiology

Received: April 20, 2016  Accepted: July 28, 2016  Published: August 25, 2016

ABSTRACT

Interleukin-17 plays a crucial role in inflammation-related carcinogenesis. We hypothesize that genetic variants in \textit{IL-17} are associated with gastric cancer (GCa) risk, and we genotyped five potentially functional single nucleotide polymorphisms (SNPs) (rs1974226 G > A, rs2275913 A > G, rs3819024 A > G, rs4711998 A > G, and rs8193036 C > T) of \textit{IL-17} in 1121 GCa patients and 1216 cancer-free controls in an eastern Chinese population. Logistic regression analysis was used to calculate odds ratios (OR) and 95% confidence intervals (CI). Meta-analysis and genotype-mRNA expression correlation were performed to further validate positive associations. We found that an increased GCa risk was independently associated with rs1974226 (adjusted OR = 2.60, 95% CI = 1.27–5.32 for AA vs. GG + GA) and rs2275913 (adjusted OR = 1.33, 95% CI = 1.03–1.72 for GA + AA vs. GG), while a decreased GCa risk was independently associated with rs3819024 (adjusted OR = 0.72, 95% CI = 0.54–0.96 for GG vs. AA + AG). Additional meta-analyses confirmed the observed risk association with rs2275913. We also found that two \textit{IL-17} haplotypes (G-G-G-A-C) and (A-G-G-A-C) (in the order of rs1974226, rs2275913, rs3819024, rs4711998 and rs8193036) were associated with a reduced GCa risk (adjusted OR = 0.64, 95% CI = 0.46–0.89 and adjusted OR = 0.38, 95% CI = 0.17–0.81, respectively). However, the expression Quantitative Trait Locus (eQTL) analysis for the genotype-phenotype correlation did not find mRNA expression changes associated with either the genotypes. In conclusions, genetic variants of \textit{IL-17} are likely to be associated with risk of GCa, and additional larger studies with functional validation are needed to explore the molecular mechanisms underlying the observed associations.
INTRODUCTION

Despite the advances in its diagnosis and therapy, gastric cancer (GCa) remains a great burden worldwide as the fifth most diagnosed cancer and the third most common cause for cancer-related deaths [1]. In China, there were an estimated 679,100 new cases and 498,000 deaths of GCa in 2015, ranking the second in both incidence and mortality of cancers [2]. Although the development of GCa is closely associated with personal lifestyles, such as smoking consumption, alcohol drinking and diet habit [3], the majority of GCa cases is attributable to the infection of Helicobacter pylori (H. pylori) and some to Epstein-Barr virus (EBV) [4, 5]. However, only a fraction of individuals who were exposed to these environmental factors eventually developed GCa, suggesting that genetic predisposition may have played an essential role in gastric carcinogenesis.

Interleukin-17A (IL-17A), also named IL-17, located on chromosome 6p12, encodes a glycoprotein of 155 amino acids [6]. IL-17 is the founding member of the cytokine family with six members (IL-17A, IL-17B, IL-17C, IL-17D, IL-17E, and IL-17F), it is mainly produced by the activated T cells, such as the T-helper 17 (Th17) cells, natural killer cells, mast cells and neutrophils, and it plays a crucial role in human immune system [7]. For example, IL-17 participates in mucosal immunity and fights against various bacteria, fungi, and EBV virus [8, 9]. IL-17 transcripts have been observed to be higher in mucosal biopsies of H. pylori-infected gastritis patients, particularly for those with chronic inflammation [10]. IL-17 has been also reported to reduce H. pylori infection in mice contracted with H. pylori-related gastritis [11]. As infection-induced chronic gastritis is a major factor in histological development of GCa [12, 13], IL-17 has been speculated to play an important role in gastric pathogenesis.

In fact, IL-17 is also believed to be involved in both gastric carcinogenesis and progression, but the exact mechanisms remain to be elucidated. For example, it has been reported that IL-17, at the tumorigenesis stage, acts as a tumor suppressor involved in recruitment of T lymphocytes and dendritic cells (DCs), promotion of NK cell activity, activation of cytotoxic T lymphocytes (CTLs) and inhibition of T reg repertoire in the tumor milieu [14, 15], whereas in the tumor progression stage, IL-17 inclines to have an oncogenic effect in inhibiting tumor cell apoptosis, impairing antitumor responses and promoting tumor angiogenesis, tumor metastasis and invasion [14, 15]. In GCa, some studies found that high expression levels of IL-17 predicted a better overall survival [16, 17], but others observed the opposite results [18–20].

Genetic variants, mostly represented by single nucleotide polymorphisms (SNPs), are believed to be involved in cancer susceptibility. Many studies of IL-17 SNPs and susceptibility have focused on inflammation-related diseases [21–24], because IL-17 has been assumed as an important pro-inflammatory factor. Only a few studies have investigated the role of IL-17 SNPs in susceptibility to cancers of the breast [25] and liver [26], and studies of GCa mainly focused on rs2275913 but generated conflicting conclusions [27–32].

To further test the hypothesis that there are associations between genetic variants of IL-17 and GCa risk, we genotyped the five most representative, potentially functional SNPs of rs1974226, rs2275913, rs3819024, rs4711998, and rs8193036 in the gene in a hospital-based case-control study in an eastern Chinese population of 1126 cases and 1221 cancer-free controls. We not only analyzed the associations between these SNPs and GCa risk, but also evaluated the effects of the haplotypes on GCa risk as well as possible functional mechanisms that may underlie the observed positive associations.

RESULTS

Characteristics of the study population

The present study included 1121 cases and 1216 controls in the final analysis, including 305 (27.2%) gastric cardi adenocarcinoma (GCA) and 816 (72.8%) gastric non-cardia adenocarcinoma (GNCA) as described elsewhere [33]. As shown in Supplementary Table S1, there were no statistically significant differences in distribution of age and sex between cases and controls (P = 0.501 for age and P = 0.326 for sex). Compared with the cases, the controls had more smokers (49.1% versus 38.7%; P < 0.001) and drinkers (28.5% versus 23.7%; P = 0.008), but these variables were further adjusted for in the subsequent multivariate logistic regression analyses to evaluate the independent main effects of the selected SNPs under investigation.

Associations between IL-17 SNPs and GCa risk

Genotype frequency distributions of the five selected potentially functional SNPs are presented in Table 1. All these SNPs in the controls were in accordance with those expected from the Hardy-Weinberg Equilibrium (HWE, with P > 0.05 for all). The risk estimates of each SNP in different genetic models were adjusted by age, sex, smoking status and drinking status and other four SNPs. From the results and trend tests presented in Table 1, the number of the rs2275913 A allele was significantly associated with an increased GCa risk (P = 0.013) but the number of the rs3819024 G allele was significantly associated with a decreased GC risk (P = 0.012) in an allele dose-response manner. The number of variant alleles of other three SNPs did not have such an obvious trend in association with GCa risk.

In a recessive genetic model with multivariate adjustment (Table 1), GCa risk was independently
Table 1: Logistic regression analysis of associations between selected SNPs of IL-17 and gastric cancer risk in an eastern Chinese population

<table>
<thead>
<tr>
<th>Variants</th>
<th>Genotype</th>
<th>Cases No. (%)</th>
<th>Controls No. (%)</th>
<th>p</th>
<th>Crude OR (95% CI)</th>
<th>p</th>
<th>Adjusted OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>All subjects, No.</td>
<td>GG</td>
<td>984 (87.8)</td>
<td>1049 (86.3)</td>
<td>0.002*</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IL-17A rs1974226</strong></td>
<td>GA</td>
<td>110 (9.8)</td>
<td>156 (12.8)</td>
<td>0.78 (0.60–0.97)</td>
<td>0.031</td>
<td>0.78 (0.60–1.02)</td>
<td>0.074</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>27 (2.4)</td>
<td>11 (0.9)</td>
<td>2.62 (1.29–5.30)</td>
<td>0.008</td>
<td>2.65 (1.28–5.47)</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td><strong>Trend test</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.996</td>
<td>0.817</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant</strong></td>
<td>GA + AA</td>
<td>137 (12.2)</td>
<td>167 (13.7)</td>
<td>0.278*</td>
<td>0.88 (0.69–1.12)</td>
<td>0.279</td>
<td>0.91 (0.71–1.17)</td>
<td>0.466</td>
</tr>
<tr>
<td><strong>Recessive</strong></td>
<td>GG + GA</td>
<td>1094 (97.6)</td>
<td>1205 (99.1)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IL-17A rs2275913</strong></td>
<td>AA</td>
<td>27 (2.4)</td>
<td>11 (0.9)</td>
<td>0.004*</td>
<td>2.70 (1.33–5.47)</td>
<td>0.006</td>
<td>2.60 (1.27–5.32)</td>
<td>0.009</td>
</tr>
<tr>
<td><strong>Trend test</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.996</td>
<td>0.817</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant</strong></td>
<td>GA + AA</td>
<td>791 (70.6)</td>
<td>827 (68.0)</td>
<td>0.182*</td>
<td>1.13 (0.95–1.34)</td>
<td>0.182</td>
<td>1.33 (1.03–1.72)</td>
<td>0.031</td>
</tr>
<tr>
<td><strong>Recessive</strong></td>
<td>GG + GA</td>
<td>890 (79.4)</td>
<td>965 (79.4)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IL-17A rs3819024</strong></td>
<td>AA</td>
<td>290 (25.9)</td>
<td>310 (25.5)</td>
<td>0.87 (0.69–1.10)</td>
<td>0.244</td>
<td>0.63 (0.43–0.92)</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td><strong>Trend test</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.265</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant</strong></td>
<td>AG + GG</td>
<td>831 (74.1)</td>
<td>906 (74.5)</td>
<td>0.835*</td>
<td>0.98 (0.81–1.18)</td>
<td>0.835</td>
<td>0.82 (0.63–1.08)</td>
<td>0.815</td>
</tr>
<tr>
<td><strong>Recessive</strong></td>
<td>AA + AG</td>
<td>880 (78.5)</td>
<td>920 (75.7)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IL-17A rs4711998</strong></td>
<td>AA</td>
<td>571 (50.9)</td>
<td>644 (53.0)</td>
<td>0.420*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trend test</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.584</td>
<td>0.661</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant</strong></td>
<td>AG + GG</td>
<td>550 (49.1)</td>
<td>572 (47.0)</td>
<td>0.328*</td>
<td>1.08 (0.92–1.28)</td>
<td>0.328</td>
<td>1.09 (0.92–1.30)</td>
<td>0.305</td>
</tr>
<tr>
<td><strong>Recessive</strong></td>
<td>AA + AG</td>
<td>1042 (92.9)</td>
<td>1123 (92.3)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IL-17A rs8193036</strong></td>
<td>CC</td>
<td>558 (49.8)</td>
<td>625 (51.4)</td>
<td>0.272*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trend test</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.886</td>
<td>0.839</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant</strong></td>
<td>CT + TT</td>
<td>563 (50.2)</td>
<td>591 (48.6)</td>
<td>0.434*</td>
<td>1.07 (0.91–1.26)</td>
<td>0.434</td>
<td>1.07 (0.89–1.27)</td>
<td>0.487</td>
</tr>
<tr>
<td><strong>Recessive</strong></td>
<td>CC + CT</td>
<td>1040 (92.8)</td>
<td>1113 (91.5)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TT</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.264*</td>
<td>0.84 (0.62–1.14)</td>
<td>0.265</td>
<td>0.81 (0.59–1.12)</td>
<td>0.201</td>
</tr>
</tbody>
</table>

Abbreviation: SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval; HWE, Hardy Weinberg Equilibrium.

*Chi square test for genotype distributions between cases and controls.

*Adjustment without (crude) and with age, sex, smoking and drinking status and all these five SNPs in logistic regression models.

*For additive genetic models.

*For dominant genetic models.

*For recessive genetic models.
associated with the rs1974226 AA genotype (adjusted OR = 2.60, 95% CI = 1.27–5.32 for AA vs. GG + GA) and the rs3819024 GG genotype (adjusted OR = 0.72, 95% CI = 0.54–0.96 for GG vs. AA + AG), while the rs2275913 GA + AA genotypes were independently associated with an increased risk in a dominant genetic model (adjusted OR = 1.33, 95% CI = 1.03–1.72 for GA + AA vs. GG). However, these risks were not observed for other two SNPs (Table 1).

SNP-SNP and SNP-environment interactions

As shown in Table 2, the logistic regression analyses identified a significant interaction between rs2275913 and rs3819024 SNPs (P = 0.009). However, we did not find evidence for any other significant SNP-SNP or SNP-environment interactions.

Meta-analysis for the association between rs2275913 and GCa risk

Because the rs2275913 SNP has been mostly investigated for its association with GCa risk in published studies, we were able to conduct a mini-meta-analysis with all the published data available for Asians. By including our new dataset, the meta-analysis consisted of a total of 7277 cases and 8519 controls. We found that the A allele was significantly associated with an increased GCa risk in either dominant (OR = 1.20, 95% CI = 1.12–1.28 for GA + AA vs. GG) or recessive (OR = 1.35, 95% CI = 1.11–1.63, for AA vs. GA + GG) genetic models (Figure 1A–1B). Furthermore, the shapes of Begg’s and Egger’s plots showed no significance publication bias for this mini-meta analysis (Supplementary Figure S1A–S1D). Because there were few studies about the other SNPs in IL-17, we were not able to perform the meta-analyses accordingly.

Stratification analysis

We further performed a stratification analysis by dichotomized variables of age, sex, smoking status and drinking status for the three significantly independent SNPs (i.e., rs1974226, rs2259713 and rs3819024) (Table 3). A significantly increased GCa risk was associated with the IL17 rs1974226 AA variant genotype in subgroups of older age (adjusted OR = 2.99, 95% CI = 1.12–7.95), males (adjusted OR = 2.56, 95% CI = 1.11–5.90), never-smokers (adjusted OR = 2.72, 95% CI = 1.07–6.96), never drinkers (adjusted OR = 2.87, 95% CI = 1.20–6.91). We also found that the rs3819024 GG variant genotype was associated with a decreased GCa risk in never-smoker subgroup (adjusted OR = 0.76, 95% CI = 0.58–0.98). However, the homogeneity test did not show that there were differences in stratum ORs for these dichotomized variables. These significant risk associations were not found for other SNPs in the stratification analysis (data not shown).

Association between haplotypes of IL-17 SNPs and GCa risk

We observed an increased GCa risk associated with rs1974226 in a recessive genetic model and rs2275913 in a dominant genetic model, while the rs3819024 was associated with a decreased GCa risk in a recessive genetic model. To account for all five SNPs in the same gene, we performed a haplotype analysis to explore their combined genetic effects. As shown in Table 4, we found that haplotypes G-G-G-A-C and A-G-G-A-C (in the order of rs1974226, rs2275913, rs3819024, rs4711998 and rs8193036) were significantly associated with a decreased GCa risk (adjusted OR = 0.64, 95% CI = 0.46–0.89 and adjusted OR = 0.38, 95% CI = 0.17–0.81, respectively), compared with the most frequent haplotype G-A-G-A-C.

Functional exploration with the expression
Quantitative Trait Locus (eQTL) analysis by IL-17 genotypes in lymphoblastoid cell lines

To explore the mechanisms underlying the observed associations, we further performed genotype-phenotype correlation analysis by using mRNA expression data of the lymphoblastoid cell lines derived from 79 unrelated Chinese people available in the HapMap 3 database. As shown in Supplementary Figure S2A–S2E, for each of the five SNPs (rs1974226, rs2275913, rs3819024, rs4711998 and rs8193036), none of their mRNA expression levels of IL-17 was correlated with the variant allele (P-trend = 0.491, 0.441, 0.848, 0.680, and 0.667, respectively), compared with their common alleles, respectively.

DISCUSSION

Mounting evidence has demonstrated that inflammation in tumor microenvironment has a role in promoting cancer cell proliferation and migration, while initiating and maintaining protective antitumor immunity, but the beneficial or detrimental effects of inflammation in terms of tumor development depend on the nature of inflammatory response and the tissue specificity [34, 35]. IL-17 is an effector cytokine that is produced by cells of both the innate and adaptive immune systems, constructing a bridge in the inflammatory reactions [36]. IL-17 was first discovered for its effects on synoviocytes from patients with rheumatoid arthritis [37], and later it was found to be involved in eliminating pathogens like gram-positive Propionibacterium acne, gram-negative Citrobacter rodentium, and fungi such as Candida albicans [38]. With the discovery of Th17 cell, IL-17 was identified as an important component of the tumor-associated immune response in the tumor microenvironment, but its exact role in carcinogenesis is still not fully understood.

To our best knowledge, the present study is the first to explore the role of all potentially functional SNPs
of *IL-17*, particularly rs1974226 in the 3’UTR region with a potential influence on miRNA binding, in cancer risk, because the *IL-17* rs1974226 AA variant genotype was associated with an increased GCa risk. Although no studies have reported an association between rs1974226 and cancer risk, two previous studies have investigated associations between rs1974226 and inflammation-related diseases. In a study of ulcerative colitis, the rs1974226 G allele in a haplotype was found to be protective for Crohn’s disease [22], suggesting that the A allele was relatively a risk factor, which is consistent with the findings in the present study.

Many studied investigated *IL-17* rs2275913 and GCa cancer risk but had generated conflicting conclusions, some reporting no association [27–29] and some demonstrating an association with GCa risk [30–32]. In the present study, we found that the *IL-17* rs2275913 A allele was associated with an increased GCa risk. Considering the varying ethnic difference in cancer susceptibility, we did a meta-analysis with previously published studies in Asian populations. Overall, we confirmed an association between the rs2275913 A allele and GCa risk in both dominant and recessive genetic models, which did not agree with some previously published meta-analyses [39–41], likely due to the fact that the present study added a large sample size to the present meta-analysis, resulting in an increased statistical power.

In addition, we also found that the GG genotype of rs3819024 was associated with a decreased GCa risk. There were only two published studies about the association between rs3819024 and cancer risks. One breast cancer study found that the A allele was associated with a reduced risk, if women consumed a high folate diet [25], while the other GCa study identified the rs3819024 G allele to be marginally associated with a decreased GCa risk in a recessive genetic model [29], a finding consistent with the present study. The opposite effect of the A allele in breast cancer and GCa may be disease-specific, but this speculation needs further validation by additional studies.

There was only one publication about the rs4711998 SNP and gastrointestinal cancer risk, suggesting that the A allele tended to be associated a non-significantly decreased hepatocellular carcinoma risk [26]. A breast cancer study revealed that carriers with the rs8193036 CC genotype had higher breast cancer mortality [25], but no published studies have evaluated rs8193036 for its association with cancer risk.

We assessed SNP-SNP and SNP-environment interactions for these three positive associated SNPs. We found that there was a SNP-SNP interaction between rs2275913 and rs3819024, suggesting that these two SNPs had a joint effect in modulating the risk of GCa.

In the haplotype analysis, we found that haplotypes G-G-G-G-A and A-G-G-A-C (alleles in the order of rs1974226G/A, rs2275913G/A, rs3819024A/G, rs4711998A/G, and rs8193036C/T) were associated with a decreased GCa risk. Except for the first allele in the order of variant alleles, the other four alleles in these two haplotypes were identical. In the trend analysis for the four SNPs, the rs2275913 G allele and rs3819024 G alleles were significantly associated with a decreased GCa risk, while the rs4711998A and rs8193036C alleles were insignificantly associated with a reduced GCa risk. The trend analysis results for these SNPs almost fit with their roles in the haplotypes. For rs1974226, although the AA genotype was obviously associated with a decreased GCa risk, it was not statistically significant in the trend analysis for the A allele, because it fit well with a recessive genetic model.

Since three (rs1974226, rs2275913 and rs3819024) of the five SNPs were found to be independently associated with GCa risk in the present study, we explored their molecular mechanism underlying the observed associations. Unfortunately, mRNA expression data from 79 unrelated Chinese people available in the HapMap 3 did not show the genotype-phenotype correlation. It is likely that these common variants may have limited and weak influence on mRNA expression not detectable in a sample of 79 individuals. Among the three positive SNPs, two are located at the 5’ near gene and one is located at the microRNA binding site. Additional mechanistic studies are warranted to explore their functional relevance to the observed associations.

It was regretful that we did not find the correlation between *IL-17* mRNA expression and genotypes of the SNPs. Because these SNPs were chosen as a tag of the gene, other untagged SNPs of the genes may be functional and could have been missed by this study. Future fine mapping of the gene is necessary to unravel such functional SNPs.

### Table 2: SNP-SNP and SNP-environment interactions (P value) between the three positive associated SNPs in *IL-17* (logistic regression)

<table>
<thead>
<tr>
<th>SNP-SNP interaction</th>
<th>rs1974226</th>
<th>rs2275913</th>
<th>rs3819024</th>
<th>Smoking</th>
<th>Drinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1974226</td>
<td>0.641</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2275913</td>
<td>0.398</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>0.960</td>
<td>0.474</td>
<td>0.986</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinking</td>
<td>0.390</td>
<td>0.943</td>
<td>0.112</td>
<td>0.244</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Forest plot of gastric cancer risk associated with *IL-17* rs2275913 from a meta-analysis of 15 case-control studies. The OR and 95% CI of each study are plotted with a box and a horizontal line. Quadrangles represent pooled ORs and 95% CI. (A) GA + AA vs. GG in a dominant genetic model and (B) AA vs. GG + GA in a recessive genetic model.
There were some limitations in the present study. First, although age, sex, smoking history, and drinking history were considered, there were some factors that may have contributed to GCa risk but were not included in the analyses, e.g. H. pylori infection status, family history, diet habit, nutrition status and socioeconomic status. Second, the numbers of SNPs genotyped in the present study were limited, and some rare potentially functional variants in the genes might have been missed, which may need additional sequencing. Third, we have no plasma...
and tissue samples available to do the further functional analysis. Therefore, the results of the present study should be interpreted with caution.

In conclusions, in the present study, we investigated associations between five potentially functional SNPs of the IL-17 gene and GCa risk in an eastern Chinese population. There are some published studies that have dealt with the role of SNPs in the IL-17 gene in GCa risk. One study included more SNPs but a small sample of only about 200 cases [29]. Another study included a large sample of more than 1,000 cases, but only studied one SNP in the IL-17 gene [27]. Most published studies included just one to three SNPs and usually with small samples of about 200–500 cases [30, 42–45]. The present study not only analyzed all the potentially functional SNPs located at the transcription factor binding site (TFBS) in the putative promoter region or at the microRNA (miRNA) binding site of IL-17 in a large sample size of more than 1,000 cases and 1,000 controls, but also did haplotype analysis. We found that the rs1974226 AA genotype was independently associated with an increased GCa risk in a recessive genetic model, but that the rs2275913 GA/AA genotypes were independently associated with an increased GCa risks in a dominant model, consistent with the findings from further meta-analysis of all published studies of this SNP. We also found that the rs3819024 GG genotype was independently associated with a decreased GCa risk in a recessive model. There was an SNP-SNP interaction between rs2275913 and rs3819024. We next found that haplotypes G-G-G-A-C and A-G-G-A-C (in the order of rs1974226, rs2275913, rs3819024, rs4711998 and rs8193036) were independently associated with a decreased GCa risk, compared with the most common haplotype G-A-G-A-C, consistent with the results of single locus analysis. Further genotype-mRNA expression correlation analysis for all these five SNPs, however, failed to find evidence for functional relevance for any of the variant genotypes. Larger, more stringently designed studies are needed to validate our findings.

MATERIALS AND METHODS

Study subjects

The final analysis consisted of 1,121 unrelated Han Chinese patients with newly diagnosed and histopathologically confirmed primary gastric adenocarcinoma from an ongoing molecular epidemiology study at Fudan University Shanghai Cancer Center (FUSCC) between 2009 and 2011 [33]. All patients came from eastern China, including Shanghai, Jiangsu, Zhenjiang and the surrounding area. Besides, 1,216 Han ethnic cancer-free controls, recruiting from Taizhou Longitudinal Study (TZL) in eastern China at the same period, were frequency matched to the cases on age (± 5 years) and sex (Supplementary Table S1). Blood samples of these GCa patients and cancer-free controls came from the tissue bank of FUSCC and the TZL study, respectively. All subjects in the study signed a written informed consent for donating their biological samples to the FUSCC tissue bank and TZL study for scientific research. This study was approved by the Institutional Review Board of FUSCC.

SNP selection and genotyping

The SNPs to be genotyped were selected from the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP) and SNPinfo (http://snpinfo.niehs.nih.gov/snpfunc.htm) using following criteria: (1) located at the TFBS in the putative promoter region; (2) located at the miRNA binding site; (3) the minor allele frequency (MAF) of at least 5% in Chinese populations; (4) with a low linkage disequilibrium (LD) with other SNPs using an r² threshold of < 0.8 as the cut-off value for each other, and (5) not included in the published genome-wide association studies (GWASs) (Table 5). As a result, only five SNPs in IL-17 satisfied the above criteria: rs1974226 located in the 3'UTR region, which may affect the miRNA binding site activity, and the other four SNPs (rs2275913, rs3819024, rs4711998 and rs8193036) located in the 5' near gene region, which may affect the TFBS activity (Table 5 and Figure 2A and 2B). The LD block of the five SNPs was shown in Figure 2C, suggesting that there was no LD between these SNPs. Taqman real-time PCR method was employed to genotype these five selected SNPs with genomic DNA extracted from the blood samples, and genotyping was performed as previously described [46], with a successful call rate of more than 99.5%. The discrepancy rate in 10% of the samples for duplication was less than 0.1%, and some samples were also randomly selected for direct sequencing to confirm the observed genotypes.

Genotype and mRNA expression data of lymphoblastoid cell lines from the HapMap3 database

We also used additional genotype data of IL-17 (http://hapmap.ncbi.nlm.nih.gov/downloads/genotypes/2010-05_phaseIII/) and mRNA expression data that are available online (http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-264/) for the genotype-phenotype correlation analysis. The genotyping data were from the HapMap phase 3 release 3 dataset consisting of about 1.6 million SNP genotypes of 692 individuals from 11 populations [47]. The mRNA expression data together with genotypes were derived from EBV-transformed B lymphoblastoid cell lines obtained from 726 individuals from 8 global populations from the HapMap3 Project [48].
Meta-analysis of the association between IL-17 variants and GCa risk using published studies

Briefly, the keywords search was performed with or without the Medical Subject Headings (MeSH) terms for: ‘interleukin-17 or IL-17’, ‘variant or polymorphism’, and ‘gastric or stomach’ and ‘cancer or neoplasm or malignancy’. Studies included in the meta-analysis must meet the following inclusion criteria: case-control designed studies; studies evaluating the association of IL-17 variants with GCa risk; studies with sufficient data for estimating genotype frequency. The main reasons for

Table 5: The selected, potentially functional SNPs as predicted by SNPinfo software

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP rs no.</th>
<th>Chromosome no.</th>
<th>Gene region</th>
<th>Functional prediction</th>
<th>Major/Minor allele</th>
<th>Minor frequency in Asians*</th>
<th>Minor frequency in CHB*</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-17</td>
<td>rs1974226</td>
<td>6</td>
<td>3'UTR</td>
<td>miRNA binding site</td>
<td>G/A</td>
<td>0.060</td>
<td>0.072</td>
</tr>
<tr>
<td>IL-17</td>
<td>rs2275913</td>
<td>6</td>
<td>5' near gene</td>
<td>TFBS</td>
<td>G/A</td>
<td>0.302</td>
<td>0.506</td>
</tr>
<tr>
<td>IL-17</td>
<td>rs3819024</td>
<td>6</td>
<td>5' near gene</td>
<td>TFBS</td>
<td>A/G</td>
<td>0.476</td>
<td>0.518</td>
</tr>
<tr>
<td>IL-17</td>
<td>rs4711998</td>
<td>6</td>
<td>5' near gene</td>
<td>TFBS</td>
<td>A/G</td>
<td>0.212</td>
<td>0.208</td>
</tr>
<tr>
<td>IL-17</td>
<td>rs8193036</td>
<td>6</td>
<td>5' near gene</td>
<td>TFBS</td>
<td>C/T</td>
<td>0.338</td>
<td>0.286</td>
</tr>
</tbody>
</table>

TFBS: transcription factor binding sites; UTR: untranslated region; CHB: Chinese Beijing, Han.
*Data from HapMap phase 3 in 137.

Figure 2: Chromosome and gene maps and locations of the potentially functional polymorphisms in the IL-17 gene as predicted by SNPinfo. (A) The chromosome structure showing the location of the IL-17 gene from UCSC browser (NCBI137/hg19); (B) The IL-17 gene structure showing the locations of the potential functional SNPs selected and studied in the present study; (C) Linkage disequilibrium (LD) blocks of IL-17 genes. The value within each diamond represents the pairwise correlation between SNPs (measured as $r^2$) defined by the upper left and the upper right sides of the diamond.
exclusion were abstract, reviews, duplicate data, of family cancers, not in English or Chinese, and not about GCa. Relevant studies were searched from PubMed and Chinese National Knowledge Infrastructure (CNKI) databases (Last Updated: March 4, 2016). The pooled ORs and 95% CIs were calculated under dominant and recessive models. χ²-based Q-test was used to check heterogeneity among the studies. Either the fixed-effects (the Mantel-Haenszel method) or random-effects (the DerSimonian and Laird method) model was chosen to calculate the pooled OR estimates according to the heterogeneity of study populations included in the meta-analysis. The Begg’s and Egger’s linear regressions were used to test the potential publication bias. Sensitivity analysis was also performed to assess the stability of the meta-analysis. All statistical tests were performed with STATA (version 12.0; Stata Corporation, College Station, TX).

**Statistical analysis**

Differences of categorical variables distribution were calculated by Pearson’s χ² test between cases and controls, and differences in continuous variables between cases and controls were evaluated by Student’s t test. The HWE was tested by a goodness-of-fit χ² test in the control group. GCa risks in both overall and stratified subgroup analyses were evaluated by ORs and 95% CIs with both univariate and multivariate logistic regression analyses. mRNA expression levels between samples of different genotypes were compared by the general linear regression analysis. Based on the observed genotypes, PROC HAPLOTYPE was used to generate individual haplotypes and frequencies to calculate ORs and 95% CIs associated with GCa risk in the logistic regression analysis. All tests were two-sided using the Statistical Analysis Software (SAS, v.9.4, SAS Institute, Cary, NC), and P < 0.05 was considered statistically significant.

**Abbreviations**

CI, confidence interval; eQTL, expression Quantitative Trait Locus; Epstein-Barr virus, EBV; GCa, gastric cancer; GCA, gastric cardia adenocarcinoma; GNCA, Gastric non-cardia adenocarcinoma; GWASs, genome-wide association studies; HWE, Hardy Weinberg equilibrium; H. pylori, Helicobacter pylori; IL-17, interleukin-17; LD, linkage disequilibrium; MAF, minor allele frequency; microRNA, miRNA; OR, odds ratios; SNP, single nucleotide polymorphism; TFBS, transcription factor binding site; T-helper 17, Th17.

**ACKNOWLEDGMENTS AND FUNDING**

This study was supported by the funds from China Recruitment Program of Global Experts at Fudan University, the Shanghai Committee of Science and Technology, China (Grant No. 12DZ2260100), Ministry of Science and Technology (2011BAI09B00) and Ministry of Health (201002007). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**CONFLICTS OF INTEREST**

The authors declared no conflicts of interest.

**REFERENCES**


