Prediction of gastric cancer risk: association between \textit{ZBTB20} genetic variance and gastric cancer risk in Chinese Han population

Fei Bai$^1$, Ke Xiao$^{1,\#}$

$^1$ Department of Gastroduodenal Pancreas Surgery, Hunan Cancer Hospital.

\textit{\# Corresponding Author:}

Ke Xiao

E-mail: kexiao_611@163.com

Tel/Fax: +86-13574840616

Address: #283 Tongzipo Road, Yuelu District, Changsha, 410013, Hunan, P.R. China
Abstract

Background: Gastric cancer (GC) is a complex multifactorial disease. Previous studies have revealed genetic variations associated with the risk of gastric cancer. The purpose of this study was to determine the correlation between single nucleotide polymorphisms (SNPs) of ZBTB20 and the risk of gastric cancer in Chinese Han population.

Methods: We conducted a 'case-control' study involving 509 GC patients and 507 healthy individuals. We selected 4 SNPs of ZBTB20 (10934270 T/C, rs9288999 G/A, rs9841504 G/C and rs73230612 C/T), and used logistic regression to analyze the relationship between those SNPs and GC risk under different genetic models; multi-factor dimensionality reduction (MDR) was used to analyze the interaction of “SNP-SNP” in gastric cancer risk; ANOVA and univariate analysis were used to analyze the differences in clinical characteristics among different genotypes.

Results: Our results showed that ZBTB20 rs9288999 is a protective factor for the risk of gastric cancer in multiple genetic models, of which the homozygous model is the most significant (OR = 0.48, p = 0.0003); we also found that rs9288999 showed a significant correlation with reducing the risk of gastric cancer in different subgroups (BMI; age; gender; smoking or drinking status; adenocarcinoma); rs9841504 is associated with increased GC risk in the participants with BMI > 24 kg/m²; rs9841504 and rs73230612 are certainly associated with clinical characteristics of platelet and carbohydrate antigen 242, respectively.

Conclusion: Our results suggest that ZBTB20 rs9288999 may be important for reducing the risk of GC in the Chinese Han population.

Keywords: Gastric cancer; ZBTB20; Polymorphisms; Case-control study
Declarations

Conflicts of interest

The authors declare that they have no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics approval and consent to participate

This study was conducted under the standard approved by Hunan Cancer Hospital, and conformed to the ethical principles for medical research involving humans of the World Medical Association Declaration of Helsinki. All participants signed informed consent forms before participating in this study.

Consent to publication

All the authors agreed to publish the manuscript.

Availability of data and materials

The datasets used and analyzed in the current study are available from the corresponding author on reasonable request.

Authors' contribution statement

The work presented here was carried out in collaboration between all authors. Juan Pan and Fei Bai carried out the molecular genetic studies and drafted the manuscript; Ke Xiao and Fei Bai designed the methods and experiments; Fei Bai performed the statistical analyses and interpreted the results. Fei Bai designed primers, performed the SNP genotyping experiments and worked on associated data collection and their interpretation. Ke Xiao and Fei Bai conceived of the study, participated in the design and coordination of the study. All authors read and approved the final manuscript.

Acknowledgments

We thank all authors for their contributions and supports. We are also grateful to all participants for providing blood samples.
Introduction

Gastric cancer (GC) is considered to be a common gastrointestinal tumor in the world, the incidence is second only to lung cancer, breast cancer and colorectal cancer, and the survival rate is low (1). The formation of gastric cancer is a complex process. As a multifactorial disease, gastric cancer is affected by both the environment and genetics (2). A number of studies have confirmed that genetic factors are important causes of gastric cancer: Mohammad et al. suggested that people who are directly related to gastric cancer patients have a higher risk of gastric cancer than normal people (3); Kaurah et al. found that 30%-40% of patients with familial diffuse gastric cancer carry the CDH1 gene mutation(4). As we all know, single nucleotide polymorphisms (SNPs) are the main form of genetic differences between individuals. With the development of molecular epidemiology and the improvement and application of genetic testing techniques, genetic polymorphisms associated with GC susceptibility have been identified in different populations (5-9).

Zinc finger and BTB domain containing 20 (ZBTB20) protein is a member of the zinc finger protein subfamily. ZBTB20 is a key regulator of α-fetoprotein expression in the adult liver. At the same time, ZBTB20 has a high degree of homology with BCL6 protein, and plays an important regulatory role in hematopoiesis, immune response, and tumor development (10, 11), and some studies have clarified the regulatory effect of ZBTB20 on gastric migration, invasion and proliferation of GC (12). Many scholars began to be interested in the association between the ZBTB20 gene single nucleotide polymorphism and gastric cancer susceptibility (13-15), but the results of these studies are not always same. More importantly, the incidence of gastric cancer is different in Eastern and Western countries, and even in different regions and different genders in China (16). The above-mentioned studies suggest that individuals with different genetic backgrounds may have different susceptibility to gastric cancer. Therefore, it is necessary to expand the scope of the study to analyze the correlation between ZBTB20 gene polymorphism and susceptibility of gastric cancer in different populations. Digging out biomarkers for gastric cancer susceptibility has potential biological and public health significance.

In summary, we conducted a correlation study between the ZBTB20 SNPs and gastric cancer susceptibility in the Chinese Han population. In this study, 509 gastric cancer patients and 507 healthy individuals were collected at the same time and place. Combined with clinical data, we
selected four sites rs10934270 T/C, rs9288999 G/A, rs9841504 G/C, rs73230612 C/T on ZBTB20 gene for our study. Then, the correlation between ZBTB20 SNPs and GC susceptibility in Chinese Han population was assessed, it will expand the association data between ZBTB20 genetic variation and susceptibility of gastric cancer.

Materials and methods

Study subjects

This study used a 'case-control' experimental design to assess the correlation between the SNPs of the gene ZBTB20 and the risk of GC in 1016 participants. These participants consisted of 509 patients with GC diagnosed at Hunan Cancer Hospital and 507 healthy individuals from the same hospital during the same period. We conducted interviews with the study subjects, and this work was carried out by professional doctors. After the interview, a complete questionnaire was formed, which included basic demographic and epidemiological information (age, gender, smoking, drinking, lymph node metastasis, pathological grade, medical history, etc.). Then, all participants provided blood samples for subsequent DNA extraction. Our study was approved by the Ethics Committee of Hunan Cancer Hospital, and all research work was carried out after the informed consents were obtained from all participants.

Selection of SNPs

After consulting the relevant literature (13, 17-22) and the data of the ZBTB20 gene polymorphism in the database, we selected SNPs with minor allele frequency ≥ 5%. Finally, 4 SNPs of ZBTB20 (rs10934270, rs9288999, rs9841504, rs73230612) were selected by us for the study.

DNA extraction and genotyping

We carried out the extraction and purification of whole genomic DNA according to the experimental steps on the kit (GoldMag Co. Ltd. Xi’an, China) instructions. Subsequently, the extracted DNA was stored in a low temperature refrigerator (-80 °C) until needed. The primers required for the study were designed by MassARRAY Assay Design software, supplemental table 1 summarized all the primers used for polymerase chain reaction (PCR) amplification and sequencing in this study. Then MassARRAY system (Agena, San Diego, CA, USA) was used for genotyping.
In order to reduce the influence of experimental operation errors on the research results and ensure the reliability and repeatability of the results of this study, we randomly selected 5% DNA samples for repeatability testing, and the experimental result repetition rate was >99%. The above steps can ensure the reliability and repeatability of the results of this study.

**Statistical analysis**

Differences in the demographic characteristics of the study were carried out using SPSS software for $\chi^2$ test, the $p$ value indicates whether it is statistically significant ($p < 0.05$: statistically significant). After the testing whether the four candidate SNPs were in Hardy-Weinberg equilibrium (SPSS software), we conducted an overall analysis and stratified analysis (age, gender, smoking or drinking status and adenocarcinoma, etc.) of the association between ZBTB20 gene polymorphism and GC risk. Using wild-type alleles as reference, plink 1.07 software was used to estimate multiple genetic models (codominant, dominant, recessive, and logarithmic addition). The results of this study are all estimated, based on the odds ratio (OR) and 95% confidence interval (CI) derived from the analysis of the logistic regression model adjusted for age and gender (OR: relative risk; OR = 1: this factor has no effect on the occurrence of disease; OR < 1: reduce the risk of disease; OR > 1: increase the risk of disease). Then, we used multi-factor dimensionality reduction (MDR) to evaluate the candidate ‘SNP-SNP’ interaction in the risk of gastric cancer. Finally, we used one-way ANOVA to predict the differences in clinical characters of gastric cancer in different genotypes. All tests in this study were two-sided tests, and $p < 0.05$ was considered statistically significant.

**Results**

**Research objects**

The 1016 participants in this study had no relationship in genetic. 'Case-control': a case group with an average age of 61.12 ± 11.33 years (male: 382 patients, proportion: 75%; female: 127 patients, proportion: 25%) and a control group with an average age of 61.35 ± 8.84 years (male: 379 healthy individuals, proportion: 75%; female: 128 healthy individuals, proportion: 25%). The results showed that there were no statistical differences in gender and age between the case group and the control group (Table 1). Table 1 summarized the demographic data of all participants, including age, gender, lymph node metastasis, adenocarcinoma, smoking and alcohol consumption,
Table 1 Characteristics of patients with gastric cancer (GC) and healthy individuals.

| Characteristics                        | Cases          | Control        | p   |
|----------------------------------------|----------------|----------------|-----|
|                                        | n = 509        | n = 507        |     |
| Age (years)                            | Mean ± SD      |                |     |
|                                        | 61.12 ± 11.33  | 61.35 ± 8.84   | 0.712 |
|                                        | > 60           |                |     |
|                                        | 279 (55%)      | 315 (62%)      |     |
|                                        | ≤ 60           |                |     |
|                                        | 230 (45%)      | 192 (38%)      |     |
| Gender                                 | Male           |                |     |
|                                        | 382 (75%)      | 379 (75%)      | 0.942 |
|                                        | Female         |                |     |
|                                        | 127 (25%)      | 128 (25%)      |     |
| Lymph node metastasis                  | Yes            |                |     |
|                                        | 235 (46%)      |                |     |
|                                        | No             |                |     |
|                                        | 97 (19%)       |                |     |
| pathological grade                     | III & IV       |                |     |
|                                        | 239 (47%)      |                |     |
|                                        | I & II         |                |     |
|                                        | 109 (21%)      |                |     |
| Adenocarcinoma                         | Yes            |                |     |
|                                        | 314 (62%)      |                |     |
| Smoking                                | Yes            |                |     |
|                                        | 233 (56%)      | 114 (22%)      |     |
|                                        | No             |                |     |
|                                        | 270 (53%)      | 172 (34%)      |     |
| Drinking                               | Yes            |                |     |
|                                        | 133 (26%)      | 119 (23%)      |     |
|                                        | No             |                |     |
|                                        | 357 (70%)      | 142 (28%)      |     |
| BMI (kg/m²)                            | BMI > 24       |                |     |
|                                        | 72 (14%)       | 183 (36%)      |     |
|                                        | BMI ≤ 24       |                |     |
|                                        | 401 (79%)      | 170 (34%)      |     |

BMI: body mass index.

Information about genotyping and candidate SNPs

Four candidate SNPs (rs10934270, rs9288999, rs9841504, rs73230612) on ZBTB20 were successfully genotyped. Detailed information about these four candidate SNPs was summarized in Table 2. All candidate SNPs were in HWE (p > 5%), and they are all located in the intron region. The results of HaploReg indicate that the candidate SNPs in this study are regulated by a variety of factors, including: SiPhy cons; DNAse; Motifs changed; Selected eQTL hits and Enhancer histone marks etc.
### Table 2 The basic information and HWE about the selected SNPs of ZBTB20.

| Gene  | SNP ID   | Role  | Chr: Position | Alleles (A/B) | MAF   | HWE (p Value) | Haploreg 4.1                                                                 |
|-------|----------|-------|---------------|---------------|-------|---------------|-----------------------------------------------------------------------------|
| ZBTB20| rs10934270| Intron| 3: 114384900  | T/C           | 0.102 | 0.095         | 0.434 SiPhy cons; DNase; Motifs changed; Selected eQTL hits                 |
| ZBTB20| rs9288999 | Intron| 3: 114429080  | G/A           | 0.357 | 0.432         | 0.928 Enhancer histone marks; Motifs changed                               |
| ZBTB20| rs9841504 | Intron| 3: 114643917  | G/C           | 0.145 | 0.134         | 0.848 Enhancer histone marks; Motifs changed; NHGRI/EBI GWAS hits           |
| ZBTB20| rs73230612| Intron| 3: 115131989  | C/T           | 0.442 | 0.416         | 0.201 Motifs changed                                                         |

HWE: Hardy–Weinberg equilibrium;

\( p > 0.05 \) indicates that the genotypes were in Hard-weinberg Equilibrium;

SNP: Single nucleotide polymorphisms;

MAF: minor allele frequency.
Correlation assessment of ZBTB20 SNPs and GC risk (overall analysis)

The correlation between SNPs and GC risk under multiple genetic models was tested based on logistic regression, and the results were corrected by age and gender. The results showed that among the 4 candidate SNPs in this study, only rs9288999 had a certain correlation with gastric cancer risk. Specifically: rs9288999 of ZBTB20 is a protective factor (OR < 1) for the risk of gastric cancer under the allele model (G vs. A, OR = 0.73, CI = 0.61 – 0.87, p = 0.001), homozygous model (GG vs. AA, OR = 0.48, CI = 0.33 – 0.71, p = 0.0003), dominant model (GG-GA vs. AA, OR = 0.72, CI = 0.56 – 0.94, p = 0.014), recessive model (GG vs. GA-AA, OR = 0.55, CI = 0.38 – 0.78, p = 0.001), and additive model (OR = 0.72, CI = 0.60 – 0.87, p = 0.0005).

There is no correlation between the remaining three candidate SNPs and the risk of gastric cancer. The above results are summarized in Table 3.

Table 3 Analysis of the association between susceptibility of gastric cancer and single nucleotide polymorphism of ZBTB20.

| SNP ID   | Model   | Genotype | Case | Control | Adjusted by age and gender |
|----------|---------|----------|------|---------|----------------------------|
|          |         |          |      |         | OR (95% CI) | p               |
| rs10934270 | Allele  | T        | 104  | 96      | 1.09 (0.81 - 1.46) | 0.571 |
|          |         | C        | 914  | 918     | 1.00           |      |
|          | Genotype| TT       | 7    | 6       | 1.17 (0.39 - 3.52) | 0.777 |
|          |         | TC       | 90   | 84      | 1.08 (0.78 - 1.50) | 0.629 |
|          |         | CC       | 412  | 417     | 1.00           |      |
|          | Dominant| TT-TT    | 97   | 90      | 1.09 (0.79 - 1.50) | 0.595 |
|          |         | CC       | 412  | 417     | 1.00           |      |
|          | Recessive| TT      | 7    | 6       | 1.16 (0.39 - 3.47) | 0.796 |
|          |         | TC-CC    | 502  | 501     | 1.00           |      |
|          | Log-additive| -       | -    | -       | 1.08 (0.81 - 1.45) | 0.583 |
| rs9288999 | Allele  | G        | 363  | 438     | 0.73 (0.61 - 0.87) | **0.001*** |
|          |         | A        | 653  | 576     | 1.00           |      |
|          | Genotype| GG       | 57   | 95      | 0.48 (0.33 - 0.71) | **0.003*** |
|          |         | GA       | 249  | 248     | 0.82 (0.62 - 1.07) | 0.140 |
|          |         | AA       | 202  | 164     | 1.00           |      |
|          | Dominant| GG-GA    | 306  | 343     | 0.72 (0.56 - 0.94) | **0.014*** |
|          |         | AA       | 202  | 164     | 1.00           |      |
|          | Recessive| GG     | 57   | 95      | 0.55 (0.38 - 0.78) | **0.001*** |
|          |         | GA-AA    | 451  | 412     | 1.00           |      |
|          | Log-additive| -       | -    | -       | 0.72 (0.60 - 0.87) | **0.0005*** |
| rs9841504 | Allele  | G        | 148  | 136     | 1.10 (0.85 - 1.41) | 0.464 |
|          |         | C        | 870  | 878     | 1.00           |      |
Correlation assessment of ZBTB20 SNPs and GC risk (subgroup analysis)

**Age or gender:** In participants ≤ 60 years old, rs9288999 of ZBTB20 may reduce the risk of GC in multiple genetic models (allele model: OR = 0.59, \( p = 0.0003 \); homozygous model: OR = 0.33, \( p = 0.0002 \); dominant model: OR = 0.60, \( p = 0.017 \), recessive model: OR = 0.39, \( p = 0.0003 \) and log-additive model: OR = 0.60, \( p = 0.0003 \)). While among the male participants, the results showed that rs9288999 was significantly associated with reduction of gastric cancer risk in multiple genetic models (allele model: OR = 0.70, \( p = 0.001 \); homozygous model: OR = 0.45, \( p < 0.001 \); dominant model: OR = 0.68, \( p = 0.012 \), recessive model: OR = 0.52, \( p = 0.002 \) and log-additive model: OR = 0.69, \( p = 0.001 \)). The specific information is summarized in Table 4.

**Smoking or drinking:** The results show that rs9288999 was associated with reducing the risk of gastric cancer among participants who do not smoke or drink alcohol (OR < 1, \( p < 0.05 \)) in homozygous (GG vs. AA) and recessive models (GG vs. GA-AA). In this subgroup analysis, there was no association between the remaining three candidate SNPs and risk of GC. The specific information is shown in Table 5.
The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (age and gender).

| SNP ID     | Model       | genotype | Age, years | Gender |
|------------|-------------|----------|------------|--------|
|            |             |          | OR (95% CI) | p      | OR (95% CI) | p      | OR (95% CI) | p |
| rs10934270 | Allele T    |          |            |        |            |        |            |    |
|            |             | C        | 1.24(0.84-1.81) | 0.280  | 0.91(0.58-1.43) | 0.681  | 1.01(0.55-1.85) | 0.978 |
|            | Genotype TT |           | 0.82(0.15-4.58) | 0.820  | 2.10(0.40-11.11) | 0.384  | 1.01(0.06-16.46) | 0.993 |
|            |             | TC       | 1.43(0.93-2.21) | 0.106  | 0.75(0.44-1.25) | 0.269  | 1.01(0.52-1.96) | 0.979 |
|            |             | CC       | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
|            | Dominant TT-TC |       | 1.39(0.91-2.12) | 0.130  | 0.82(0.50-1.34) | 0.422  | 1.01(0.53-1.93) | 0.978 |
|            |             | CC       | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
|            | Recessive TT |           | 0.77(0.14-4.28) | 0.763  | 2.2(0.42-11.63) | 0.353  | 1.01(0.06-16.40) | 0.994 |
|            |             | TC-CC    | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
| rs9288999  | Allele G    |          | 0.83(0.65-1.05) | 0.116  | 0.59(0.45-0.78) | 0.0003* | 0.82(0.57-1.17) | 0.271 |
|            |             | A        | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
|            | Genotype GG |           | 0.65(0.38-1.12) | 0.122  | 0.33(0.18-0.59) | 0.0002* | 0.61(0.29-1.30) | 0.199 |
|            |             | GA       | 0.83(0.58-1.18) | 0.298  | 0.75(0.48-1.17) | 0.207  | 0.96(0.56-1.65) | 0.881 |
|            |             | AA       | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
|            | Dominant GG-GA |        | 0.79(0.56-1.11) | 0.171  | 0.60(0.40-0.91) | 0.017* | 0.86(0.51-1.44) | 0.567 |
|            |             | AA       | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
|            | Recessive GG |           | 0.73(0.44-1.20) | 0.213  | 0.39(0.23-0.65) | 0.0003* | 0.62(0.31-1.24) | 0.179 |
|            |             | GA-AA    | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
| rs9841504  | Allele G    |          | 1.04(0.75-1.43) | 0.820  | 1.26(0.83-1.90) | 0.276  | 1.22(0.74-2.00) | 0.429 |
|            |             | C        | 1.00        | 1.00    | 1.00        | 1.00    | 1.00        | 1.00  |
| SNP: Single nucleotide polymorphisms; | OR, Odds ratio; | CI, Confidence interval; | p < 0.05 indicates statistical significance; | "-" indicates Log-additive model. |
|--------------------------------------|----------------|----------------------|---------------------------------|----------------------------------|
| Genotype                             |                |                      |                                 |                                  |
| GG                                   | 1.16(0.38-3.57) | 0.797                | 6.47(0.79-52.99)              | 0.082 | 2.62(0.49-13.93) | 0.259 | 1.50(0.53-4.26) | 0.450 |
| GC                                   | 1.11(0.76-1.62) | 0.587                | 1.00(0.62-1.61)              | 0.993 | 1.04(0.58-1.87) | 0.890 | 1.00(0.72-1.40) | 0.998 |
| CC                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Dominant                             |                |                      |                                 |                                  |
| GG-GC                                | 1.12(0.77-1.61) | 0.563                | 1.13(0.71-1.80)              | 0.602 | 1.14(0.65-2.00) | 0.640 | 1.03(0.74-1.43) | 0.852 |
| GC                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| CC                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Recessive                            |                |                      |                                 |                                  |
| GG                                   | 1.13(0.37-3.45) | 0.835                | 6.48(0.79-52.89)              | 0.081 | 2.60(0.49-13.74) | 0.262 | 1.50(0.53-4.25) | 0.449 |
| GC                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| CC                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Log-additive                         |                |                      |                                 |                                  |
| rs73230612                            | 1.10(0.79-1.53) | 0.566                | 1.25(0.83-1.88)              | 0.294 | 1.21(0.75-1.96) | 0.441 | 1.06(0.79-1.42) | 0.701 |
| Allele                               |                |                      |                                 |                                  |
| C                                    | 1.12(0.89-1.41) | 0.331                | 1.09(0.83-1.43)              | 0.542 | 0.86(0.61-1.23) | 0.412 | 1.21(0.99-1.48) | 0.068 |
| T                                    | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Genotype                             |                |                      |                                 |                                  |
| CC                                   | 1.09(0.68-1.75) | 0.714                | 1.17(0.66-2.05)              | 0.597 | 0.75(0.37-1.53) | 0.435 | 1.38(0.92-2.08) | 0.123 |
| CT                                   | 1.28(0.88-1.86) | 0.190                | 1.19(0.77-1.85)              | 0.436 | 0.86(0.50-1.50) | 0.599 | 1.42(1.03-1.96) | 0.035 |
| TT                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Dominant                             |                |                      |                                 |                                  |
| CC-CT                                | 1.22(0.86-1.73) | 0.256                | 1.18(0.78-1.80)              | 0.427 | 0.83(0.49-1.39) | 0.481 | 1.41(1.04-1.91) | 0.586 |
| TT                                   | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Recessive                            |                |                      |                                 |                                  |
| CC                                   | 0.95(0.62-1.44) | 0.801                | 1.05(0.64-1.72)              | 0.858 | 0.82(0.44-1.54) | 0.543 | 1.13(0.78-1.62) | 0.526 |
| CT-TT                                | 1.00            | 1.00                 | 1.00                          | 1.00                                      |
| Log-additive                         |                |                      |                                 |                                  |
| -                                    | 1.08(0.85-1.35) | 0.537                | 1.09(0.83-1.45)              | 0.528 | 0.87(0.61-1.23) | 0.422 | 1.21(0.98-1.47) | 0.070 |

The content refers to genetic data, with annotations indicating different genotypes and models used for analysis. The table includes odds ratios and confidence intervals for various genotypic comparisons, highlighting significant p-values for statistical significance.
Table 5 The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (smoking and drinking).

| SNP ID    | Model | genotype | Smoking  |  |  |  | Drinking |
|-----------|-------|----------|----------|  |  |  |----------|
|           |       |          | OR (95% CI) | p | OR (95% CI) | p | OR (95% CI) | p |
| rs10934270| Allele| T        | 0.90(0.56-1.46) | 0.666 | 0.86(0.54-1.36) | 0.523 | 0.91(0.54-1.56) | 0.744 |
|           |       | C        | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Genotype | TT     | 1.30(0.26-6.59) | 0.754 | 0.29(0.03-3.25) | 0.315 | 2.04(0.39-10.85) | 0.401 |
|           |       | TC      | 0.76(0.43-1.33) | 0.335 | 0.93(0.56-1.53) | 0.763 | 0.72(0.37-1.38) | 0.317 |
|           |       | CC      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Dominant | TT-TC | 0.80(0.46-1.38) | 0.419 | 0.89(0.54-1.45) | 0.629 | 0.82(0.44-1.51) | 0.525 |
|           |       | CC      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Recessive | TT    | 1.37(0.27-6.92) | 0.705 | 0.29(0.03-3.29) | 0.320 | 2.17(0.41-11.48) | 0.361 |
|           |       | TC      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Log-additive | -  | 0.87(0.55-1.39) | 0.565 | 0.85(0.54-1.36) | 0.505 | 0.95(0.57-1.58) | 0.829 |
| rs9288999 | Allele| G        | 0.78(0.56-1.08) | 0.132 | 0.77(0.58-1.01) | 0.059 | 0.75(0.52-1.08) | 0.120 |
|           |       | A        | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Genotype | GG    | 0.54(0.27-1.07) | 0.077 | 0.46(0.26-0.82) | 0.009* | 0.49(0.23-1.07) | 0.072 |
|           |       | GA      | 0.84(0.51-1.39) | 0.498 | 1.16(0.76-1.78) | 0.500 | 1.04(0.60-1.8) | 0.904 |
|           |       | AA      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Dominant | GG-GA | 0.75(0.47-1.21) | 0.237 | 0.92(0.62-1.38) | 0.690 | 0.86(0.51-1.43) | 0.555 |
|           |       | AA      | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Recessive | GG    | 0.59(0.32-1.11) | 0.101 | 0.42(0.25-0.72) | 0.002* | 0.48(0.24-0.99) | 0.047 |
|           |       | GA-AA   | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
|           | Log-additive | -  | 0.76(0.54-1.05) | 0.095 | 0.76(0.57-1.00) | 0.052 | 0.77(0.53-1.10) | 0.150 |
| rs9841504 | Allele| G        | 1.49(0.88-2.51) | 0.137 | 1.02(0.70-1.47) | 0.928 | 1.20(0.69-2.06) | 0.520 |
|           |       | C        | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Genotype | GG     | 0.999 | 2.00 (0.53-7.46) | 0.305 | /     | 0.999 | 1.48 (0.41-5.39) | 0.551 |
|----------|--------|-------|------------------|-------|-------|-------|-----------------|-------|
| GC       | 1.43 (0.80-2.56) | 0.234 | 0.87 (0.56-1.35) | 0.542 | 1.06 (0.58-1.95) | 0.840 | 0.92 (0.59-1.46) | 0.732 |
| CC       | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| Dominant | GG-GC  | 1.53 (0.86-2.73) | 0.150 | 0.94 (0.62-1.44) | 0.778 | 1.14 (0.62-2.06) | 0.678 | 0.97 (0.62-1.50) | 0.883 |
| CC       | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| recessive| GG     | /     | 0.999            | 2.07 (0.56-7.69) | 0.279 | /     | 0.999 | 1.51 (0.42-5.47) | 0.530 |
| GC-GC    | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| rs73230612| Allele |        |                  |       |       |       |                 |       |
| C        | 1.09 (0.80-1.5) | 0.581 | 1.15 (0.87-1.52) | 0.326 | 1.37 (0.96-1.95) | 0.080 | 1.12 (0.85-1.49) | 0.417 |
| T        | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| Genotype | CC     | 1.18 (0.62-2.33) | 0.618 | 1.21 (0.67-2.17) | 0.524 | 2.03 (0.97-4.25) | 0.060 | 1.17 (0.66-2.05) | 0.595 |
| CT       | 1.16 (0.68-1.98) | 0.590 | 1.29 (0.85-1.97) | 0.233 | 1.52 (0.83-2.77) | 0.175 | 1.38 (0.90-2.13) | 0.142 |
| TT       | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| Dominant | CC-CT  | 1.16 (0.70-1.92) | 0.552 | 1.27 (0.85-1.90) | 0.237 | 1.65 (0.93-2.93) | 0.085 | 1.32 (0.88-1.97) | 0.177 |
| TT       | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| recessive | CC    | 1.07 (0.62-1.86) | 0.799 | 1.05 (0.61-1.79) | 0.868 | 1.54 (0.83-2.88) | 0.172 | 0.98 (0.58-1.64) | 0.932 |
| CT-TT    | 1.00   | 1.00  | 1.00             | 1.00  | 1.00  | 1.00  | 1.00            |       |
| Log-additive | -      | 1.09 (0.79-1.50) | 0.600 | 1.14 (0.86-1.51) | 0.362 | 1.43 (0.99-2.07) | 0.057 | 1.13 (0.86-1.50) | 0.382 |

OR, Odds ratio;
CI, Confidence interval;

\( p < 0.05 \) indicates statistical significance.

"+" indicates Log-additive model;

"/" indicates data missing.
**BMI:** The results are shown in Table 6: The rs9288999 of ZBTB20 may reduce the risk of gastric cancer in the study population with BMI < 24 under the allele model (OR = 0.68, p = 0.004), homozygous model (OR = 0.41, p = 0.001), recessive model (OR = 0.46, p = 0.002) and additive model (OR = 0.67, p = 0.004); on the contrary, rs9841504 may increase the risk of gastric cancer in the study population with BMI > 24 under the homozygous model (OR = 11.9, p = 0.028) and recessive model (OR = 12.29, p = 0.026).

Table 6 The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (BMI).

| SNP ID   | Model       | genotype | BMI   | p         | OR (95% CI) | BMI   | p         |
|----------|-------------|----------|-------|-----------|-------------|-------|-----------|
|          |             |          | < 24  |           |             | > 24  |           |
| rs10934270 | Allele      | T/C      | 0.87(0.58-1.30) | 0.490 | 1.26(0.66-2.41) | 0.490 |
|          | Homozygote  | TT/CC    | 0.78(0.19-3.19) | 0.726 | 2.90(0.18-47.36) | 0.456 |
|          | Heterozygote| TC       | 0.83(0.52-1.32) | 0.425 | 1.17(0.56-2.42) | 0.680 |
|          | Dominant    | TT/TC/CC | 0.82(0.52-1.29) | 0.396 | 1.22(0.60-2.48) | 0.584 |
|          | Recessive   | TT/TC-CC | 0.80(0.20-3.30) | 0.762 | 2.82(0.17-45.89) | 0.467 |
|          | Log-additive| -        | 0.84(0.56-1.26) | 0.401 | 1.26(0.65-2.43) | 0.500 |
| rs9288999 | Allele      | G/A      | 0.68(0.53-0.89) | 0.004* | 0.80(0.54-1.19) | 0.265 |
|          | Homozygote  | GG/AA    | 0.41(0.24-0.71) | 0.001* | 0.57(0.24-1.39) | 0.220 |
|          | Heterozygote| GA       | 0.84(0.56-1.26) | 0.389 | 0.77(0.42-1.41) | 0.391 |
|          | Dominant    | GG-GA-AA | 0.70(0.48-1.03) | 0.072 | 0.72(0.40-1.29) | 0.269 |
|          | Recessive   | GG/GA-AA | 0.46(0.28-0.74) | 0.002* | 0.68(0.30-1.52) | 0.342 |
|          | Log-additive| -        | 0.67(0.51-0.88) | 0.004* | 0.76(0.50-1.15) | 0.198 |
| rs9841504 | Allele      | G/C      | 1.41(0.95-2.10) | 0.087 | 1.43(0.85-2.40) | 0.180 |
|          | Homozygote  | GG/CC    | 3.56(0.44-29.1) | 0.236 | 11.9(1.31-108.82) | 0.028* |
|          | Heterozygote| GC       | 1.32(0.84-2.05) | 0.227 | 0.87(0.45-1.69) | 0.688 |
|          | Dominant    | GG-GC/CC | 1.38(0.89-2.14) | 0.149 | 1.11(0.60-2.04) | 0.744 |
|          | Recessive   | GG/GC-CC | 3.35(0.41-27.33) | 0.259 | 12.29(1.36-111.20) | 0.026* |
|          | Log-additive| -        | 1.40(0.93-2.10) | 0.105 | 1.35(0.80-2.28) | 0.267 |
| rs73230612| Allele      | C/T      | 1.16(0.90-1.50) | 0.256 | 0.98(0.67-1.45) | 0.937 |
|          | Homozygote  | CC/TT    | 1.37(0.80-2.34) | 0.248 | 0.89(0.38-2.06) | 0.777 |
|          | Heterozygote| CT       | 1.13(0.75-1.70) | 0.558 | 1.39(0.74-2.61) | 0.302 |
|          | Dominant    | CC-CT/TT | 1.19(0.81-1.75) | 0.374 | 1.24(0.68-2.26) | 0.482 |
|          | Recessive   | CC-CT-TT | 1.27(0.79-2.04) | 0.319 | 0.72(0.34-1.51) | 0.385 |
|          | Log-additive| -        | 1.16(0.90-1.51) | 0.251 | 1.00(0.67-1.48) | 0.987 |

SNP: Single nucleotide polymorphisms;
OR, Odds ratio;
CI, Confidence interval;
p < 0.05 indicates statistical significance;
"*" indicates Log-additive model.
**Adenocarcinoma**: In gastric cancer patients with adenocarcinoma, rs9288999 is a protective factor for the risk of gastric cancer in multiple genetic models (OR < 1, p < 0.05), such as allele model (OR = 0.72, p = 0.002), homozygous model (OR = 0.52, p = 0.004), heterozygous model (OR = 0.72, p = 0.037), dominant model (OR = 0.67, p = 0.006), recessive model (OR = 0.63, p = 0.024) and additive model (OR = 0.72, p = 0.002). The specific information is summarized in Table 7.

In addition, we also divided the gastric cancer cases in this study according to the pathological grade (I & II vs. III & IV) and whether the lymph nodes metastasized. No association have been found between candidate SNPs and gastric cancer risk (supplemental table 2).

**Table 7** The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (adenocarcinoma).

| SNP ID    | Model     | genotype | Adenocarcinoma (patients with GC) | Case | Control | OR (95% CI)   | p      |
|-----------|-----------|----------|-----------------------------------|------|---------|---------------|--------|
| rs10934270 | Allele    | T/C      |                                   | 60   | 96      | 1.01(0.72-1.42) | 0.954  |
|           | Homozygote| TT/CC    |                                   | 5    | 6       | 1.30(0.39-4.32) | 0.667  |
|           | Heterozygote| TC      |                                   | 50   | 84      | 0.95(0.65-1.40) | 0.804  |
|           | Dominant  | TT/TC/CC |                                   | 55   | 90      | 0.98(0.67-1.41) | 0.898  |
|           | Recessive | TT/TC-CC |                                   | 5    | 6       | 1.31(0.40-4.35) | 0.656  |
|           | Log-additive | -    |                                   | -    | -       | 1.00(0.72-1.39) | 0.996  |
| rs9288999 | Allele    | G/A      |                                   | 222  | 438     | 0.72(0.59-0.89) | 0.002* |
|           | Homozygote| GG/AA    |                                   | 40   | 95      | 0.52(0.34-0.81) | 0.004* |
|           | Heterozygote| GA      |                                   | 142  | 248     | 0.72(0.53-0.98) | 0.037* |
|           | Dominant  | GG-GA/AA |                                   | 182  | 343     | 0.67(0.50-0.89) | 0.006* |
|           | Recessive | GG/GA-AA |                                   | 40   | 95      | 0.63(0.42-0.94) | 0.024* |
|           | Log-additive | -    |                                   | -    | -       | 0.72(0.59-0.89) | 0.002* |
| rs9841504 | Allele    | G/C      |                                   | 93   | 136     | 1.12(0.84-1.49) | 0.427  |
|           | Homozygote| GG/CC    |                                   | 8    | 8       | 1.68(0.62-4.53) | 0.309  |
|           | Heterozygote| GC      |                                   | 77   | 120     | 1.07(0.77-1.49) | 0.697  |
|           | Dominant  | GG-GC/CC |                                   | 85   | 128     | 1.11(0.80-1.52) | 0.538  |
|           | Recessive | GG/GC-CC |                                   | 8    | 8       | 1.65(0.61-4.44) | 0.322  |
|           | Log-additive | -    |                                   | -    | -       | 1.13(0.85-1.50) | 0.405  |
| rs73230612| Allele    | C/T      |                                   | 282  | 422     | 1.15(0.94-1.41) | 0.173  |
|           | Homozygote| CC/TT    |                                   | 64   | 95      | 1.28(0.86-1.92) | 0.224  |
|           | Heterozygote| CT      |                                   | 154  | 232     | 1.26(0.91-1.73) | 0.164  |
|           | Dominant  | CC-CT/TT |                                   | 218  | 327     | 1.26(0.93-1.71) | 0.128  |
|           | Recessive | CC/CT-TT |                                   | 64   | 95      | 1.12(0.79-1.60) | 0.523  |
|           | Log-additive | -    |                                   | -    | -       | 1.15(0.94-1.40) | 0.174  |
SNP: Single nucleotide polymorphisms; 
OR, Odds ratio; 
CI, Confidence interval; 
$p < 0.05$ indicates statistical significance; 
“-” indicates Log-additive model.

MDR analysis

Subsequently, MDR analysis was used to assess the interaction of 'SNP-SNP'. Figure 1 shows the interaction between the four candidate SNPs of $ZBTB20$. The blue line indicates that these four SNPs may have a redundant role in regulating the risk of diabetes. All experimental results have been shown in Table 8. The best unit point model for predicting the risk of gastric cancer is: rs9288999 (testing accuracy = 0.521, CVC = 10/10, $p = 0.0006$); the two-site model is: rs9288999, rs73230612 (testing accuracy = 0.535, CVC = 10/10, $p = 0.0003$); the three-site model is: rs10934270, rs9288999 and rs73230612 (testing accuracy = 0.500, CVC = 7/10, $p < 0.0001$); the four-site model is: rs10934270, rs9288999, rs9841504 and rs73230612 (testing accuracy = 0.517, CVC = 10/10, $p < 0.0001$). Figure 2 shows the interaction of ‘SNP-SNP’ in different combinations of sites, among them, the light gray grid represents a low risk of gastric cancer, the darker gray grid represents a higher risk of gastric cancer, and the unfilled grid represents no data. Therefore, we can conclude that the impact of these four candidate SNPs of $ZBTB20$ on the risk of gastric cancer may be interdependent.

![Figure 1](image-url) 

**Figure 1** Dendrogram analysis of SNP-SNP interaction. The colors in the tree diagram represent synergy (yellow) or redundancy (blue).
Figure 2 Multifactor dimensionality reduction (MDR) analysis for SNPs (10934270, rs9288999, rs9841504 and rs73230612) of ZBTB20 interaction. In each box, the left bar represents cases and the right bar represents controls. The light gray boxes indicate the low risk of gastric cancer and dark gray boxes indicate the high risk, the empty boxes mean no data.

Differences in clinical indicators under different genotypes

Finally, we also evaluated the correlation between the four candidate SNPs polymorphisms (rs10934270, rs9288999, rs9841504, rs73230612) and clinical indicators of gastric cancer patients. These clinical indicators include carcinoembryonic antigen (CEA), tumor necrosis factor (TNF),
carbohydrate antigen 50 (CA50), carbohydrate antigen 19-9 (CA19-9), carbohydrate antigen 242 (CA242), white blood cells (WBC), hemoglobin (HGB) and platelet (PLT). The results showed (Table 9 and Supplemental table 3): The rs9841504 of ZBTB20 had a potential association with the content of PLT ($p = 0.048$), while the rs73230612 had a significant association with CA242 ($p = 0.005$).
Table 8 SNP-SNP interaction models analyzed by the MDR method.

| Model                                      | Training Bal. Acc | Testing Bal. Acc | OR (95% CI)          | p value | CVC |
|--------------------------------------------|-------------------|------------------|----------------------|---------|-----|
| rs9288999                                  | 0.540             | 0.521            | 1.86 (1.30-2.65)     | 0.0006  | 10/10 |
| rs9288999, rs73230612                       | 0.557             | 0.535            | 1.60 (1.24-2.05)     | 0.0003  | 10/10 |
| rs10934270, rs9288999, rs73230612            | 0.569             | 0.500            | 1.72 (1.33-2.22)     | < 0.0001| 7/10 |
| rs10934270, rs9288999, rs9841504, rs73230612 | 0.586             | 0.517            | 1.96 (1.52-2.51)     | < 0.0001| 10/10 |

MDR, multifactor dimensionality reduction;
Bal. Acc., balanced accuracy;
CVC, cross-validation consistency;
OR, odds ratio;
95% CI, 95% confidence interval.

p values were calculated using χ² tests;
p < 0.05 indicates statistical significance.
Table 9 Clinical characteristics of patients based on the genotypes of selected SNPs.

| Characteristics   | rs9841504 |        |        |   | rs73230612 |        |        |   |
|-------------------|-----------|--------|--------|---|-----------|--------|--------|---|
|                   | CC        | CG     | GG     | p | TT        | TC     | CC     | p |
| CEA               | 17.28 ± 10.59 | 16.99 ± 11.16 | 14.06 ± 3.49 | 0.637 | 17.38 ± 8.85 | 16.24 ± 9.34 | 18.94 ± 15.26 | 0.230 |
| TNF (fmol/ml)     | 0.90 ± 0.07 | 1.39 ± 4.30 | 0.88 ± 0.06 | 0.196 | 0.88 ± 0.07 | 1.12 ± 2.92 | 0.90 ± 0.07 | 0.569 |
| CA50 (U/ml)       | 7.48 ± 11.94 | 7.74 ± 12.76 | 6.05 ± 7.78 | 0.916 | 8.35 ± 13.58 | 7.02 ± 11.04 | 7.45 ± 12.01 | 0.681 |
| CA19-9 (U/ml)     | 45.57 ± 89.00 | 55.45 ± 107.75 | 20.19 ± 8.30 | 0.491 | 51.65 ± 91.31 | 44.68 ± 94.12 | 44.63 ± 90.63 | 0.827 |
| CA242 (KU/ml)     | 16.48 ± 30.97 | 14.11 ± 27.37 | 5.47 ± 7.72 | 0.493 | 23.51 ± 42.37 | 11.39 ± 19.71 | 12.87 ± 23.18 | 0.005* |
| WBC (L)           | 6.27 ± 5.92 | 7.83 ± 6.67 | 3.83 ± 2.13 | 0.293 | 8.14 ± 8.17 | 5.96 ± 5.45 | 6.16 ± 3.66 | 0.127 |
| HGB (g/L)         | 102.76 ± 22.62 | 111.02 ± 28.04 | 93.5 ± 37.03 | 0.108 | 105.38 ± 26.76 | 102.33 ± 22.16 | 108.7 ± 26.79 | 0.417 |
| PLT (L)           | 189.64 ± 94.02 | 229.40 ± 132.05 | 307.00 ± 89.10 | 0.048* | 203.16 ± 99.62 | 207.62 ± 117.15 | 187.68 ± 94.63 | 0.669 |

CEA, carcinoembryonic antigen;  
TNF, tumor necrosis factor;  
CA50, carbohydrate antigen 50;  
CA19-9, carbohydrate antigen 19-9;  
CA242, carbohydrate antigen 242;  
WBC, white blood cells;  
HGB, hemoglobin;  
PLT, platelet.
Discussion

Gastric cancer is the result of multiple gene-environment interactions, and a single gene plays a smaller role in it (23). It is estimated that only a small part of the tumor susceptibility areas/sites found in GWAS can explain the occurrence of tumors. Among them, prostate cancer can reach 15%, while in breast and colorectal cancer, only 5% and 4%, even less in gastric cancer (24). Therefore, it is still a long-term and arduous task to carry out more researches on the association of gastric cancer risk to identify more unknown susceptible areas/sites. The incidence of gastric cancer varies among different societies, East Asia, Central America, South America and Eastern Europe have higher incidence, while Africa and North America have lower incidence (1, 25). There are differences in the incidence of gastric cancer even between north and south of China (16). Wherefore, it is very necessary to conduct gastric cancer association studies in different populations. These research results are of great significance for us to understand the susceptibility mechanism of gastric cancer, to explore the pathogenesis of gastric cancer, the risk prediction and screening of high-risk groups, and to guide the individualized treatment of gastric cancer.

Some cancer-related studies have found that NF-Kb (nf-kappa b) may cause oncogenesis, which is carried out by induction of genes encoding proteins. These proteins are related to invasion, migration and inhibition of cell apoptosis (26, 27). Liu et al. found that ZBTB20 can promote the activation of NF-κB through inhibiting IκBa gene transcription or regulating protein expression (28). And there were studies showed that overexpression of ZBTB20 can promote the proliferation, migration and invasion of gastric cancer cells, which may be regulated by the IκBa/NF-κB signaling pathway (12). The above research results suggested that ZBTB20 may be a target for gastric cancer prevention and treatment. For genetics field, some studies have pointed out that genetic mutations and gene polymorphisms are the main reasons for the differences in the gastric risk among individuals (29, 30). Some polymorphism sites are significantly associated with gastric cancer risk have been reported (31-34), but there was no research about the association between ZBTB20 and the risk of gastric cancer in Chinese Han population.

Therefore, in our study, the Chinese Han population was used as the research object to conduct a ‘case-control’ study, then we analyzed the correlation between ZBTB20 gene polymorphisms (rs10934270, rs9288999, rs9841504, rs73230612) and gastric cancer risk. The
results of our study showed that ZBTB20 rs9288999 was significantly associated with reducing the risk of gastric cancer among the study population, whether in the overall analysis or stratified analysis. Dong, L. et al. and Yusefi, A. et al. proposed that genetic polymorphisms may affect disease risk by regulating the expression of certain genes (29, 30). Combining the results of our study, we speculated that the polymorphic site ‘rs9288999’ may have regulated the expression of ZBTB20 through the IκBα/NF-κB signaling pathway, which made rs9288999 showed a significant association with reduction of gastric cancer risk in the Chinese Han population. But this complicated process still needs more comprehensive research to verify. Nevertheless, as far as we know, this study is the first to find evidence of a correlation between ZBTB20 rs9288999 and gastric cancer risk. At the same time, the results of our study also provided new genetic-related research directions for gastric cancer prevention and treatment in the future.

Since gastric cancer is a multifactorial disease, the prerequisite for prevention is to accurately identify and manage the risk factors or potential causes of gastric cancer (35). Studies have found evidence that people who smoke and drink alcohol may have a higher risk of gastric cancer than those who do not smoke or drink alcohol (36-38). Lope et al. found that the case group (patients with gastric cancer) was significantly older than the control group (39). Studies have shown that gender factors may also play a role in gastric cancer risk (40). We got similar results and found that there was a significant correlation between ZBTB20 rs9288999 and reducing the gastric cancer risk in participants who are non-smokers (OR = 0.46, p = 0.009), non-drinking (OR = 0.54, p = 0.035), age ≤ 60 years (OR = 0.33, p = 0.0002) and male participants (OR = 0.45, p = 0.0005). Based on previous studies and our study, the cause of the above results may be the interaction between environmental factors and genetic polymorphisms. what’s more, ZBTB20 rs9288999 may play a certain role in that.

In addition, exploring the interaction between SNP-SNP can also help us to discover potential risk factors for the incidence of gastric cancer. Therefore, MDR was used to explore the interaction between the four candidate SNPs. The results showed that for gastric cancer risk, rs10934270, rs9288999, rs9841504 and rs73230612 showed a strong interaction.

In the correlation analysis between clinical indicators and gastric cancer risk, we also found that the ZBTB20 polymorphic site rs9841504 and platelet showed a certain significant correlation (p = 0.048). Platelet involves in cancer growth and metastasis is a long-term concept (41), and...
studies have shown that the ratio of platelet to lymphocyte and CA242 may be convenient biomarkers for gastric cancer prognosis (42, 43). Therefore, our results suggested that rs9841504 may play a certain role in the influence of platelets on the occurrence and development of gastric cancer. However, due to the limitation of sample size and ethnicity, this result can only be used as a reminder. It requires deeper research to be accurately verified.

This study provides data supplements for the association between the ZBTB20 gene polymorphisms and the risk of gastric cancer in the Chinese Han population, and concludes that there may have certain association between the two. However, we must face the fact that our research has certain limitations, which is not only for the confirmation of results but also for new discoveries, a large sample size is indeed necessary. Currently, the genetic regions/sites discovered are only a small part, and there are more genetic susceptible sites/regions waiting to be discovered. With the discovery of susceptibility sites for gastric cancer in the future, we will have a more comprehensive understanding of the genetics of gastric cancer.

**Conclusion**

In summary, our study is the first to find that the rs9288999 of ZBTB20 has a potential association with reducing the risk of gastric cancer in the Chinese Han population. It provides new data supplement for the study of the association between ZBTB20 gene polymorphism and gastric cancer risk.
Reference

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians. 2011;61(2):69-90.

2. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2014;23(5):700-13.

3. Yaghoobi M, McNabb-Baltar J, Bijarchi R, Hunt RH. What is the quantitative risk of gastric cancer in the first-degree relatives of patients? A meta-analysis. World journal of gastroenterology. 2017;23(13):2435-42.

4. Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N, et al. Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. Jama. 2007;297(21):2360-72.

5. Lu X, Chen F, Liu X, Yuan D, Zi Y, He X, et al. Detection and clinical significance of DNA repair gene ERCC8 tag SNPs in gastric cancer. The Turkish journal of gastroenterology : the official journal of Turkish Society of Gastroenterology. 2018;29(4):392-6.

6. Gonzalez-Hormazabal P, Romero S, Musleh M, Bustamante M, Stambuk J, Pisano R, et al. IL-8-251T>A (rs4073) Polymorphism Is Associated with Prognosis in Gastric Cancer Patients. Anticancer research. 2018;38(10):5703-8.

7. Zabaleta J. Multifactorial etiology of gastric cancer. Methods in molecular biology (Clifton, NJ). 2012;863:411-35.

8. Jia ZF, Cao DH, Wu YH, Jin MS, Pan YC, Cao XY, et al. Lethal-7-related polymorphisms are associated with susceptibility to and prognosis of gastric cancer. World journal of gastroenterology. 2019;25(8):1012-23.

9. Xu R, Peng C, Xiao S, Zhuang W. IFRD1 polymorphisms and gastric cancer risk in a Chinese population. Medical oncology (Northwood, London, England). 2014;31(9):135.

10. Sutherland AP, Zhang H, Zhang Y, Michaud M, Xie Z, Patti ME, et al. Zinc finger protein Zbtb20 is essential for postnatal survival and glucose homeostasis. Molecular and cellular biology. 2009;29(10):2804-15.

11. Zhang W, Mi J, Li N, Sui L, Wan T, Zhang J, et al. Identification and characterization of DPZF, a novel human BTB/POZ zinc finger protein sharing homology to BCL-6. Biochemical and biophysical
12. Zhang Y, Zhou X, Zhang M, Cheng L, Zhang Y, Wang X. ZBTB20 promotes cell migration and invasion of gastric cancer by inhibiting IkBα to induce NF-κB activation. Artificial cells, nanomedicine, and biotechnology. 2019;47(1):3862-72.

13. 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(12):1215-8.

14. Song HR, Kim HN, Kweon SS, Choi JS, Shim HJ, Cho SH, et al. Genetic variations in the PRKAA1 and ZBTB20 genes and gastric cancer susceptibility in a Korean population. Molecular carcinogenesis. 2013;52 Suppl 1:E155-60.

15. Cai M, Dai S, Chen W, Xia C, Lu L, Dai S, et al. Environmental factors, seven GWAS-identified susceptibility loci, and risk of gastric cancer and its precursors in a Chinese population. 2017;6(3):708-20.

16. Yang L. Incidence and mortality of gastric cancer in China. World journal of gastroenterology. 2006;12(1):17-20.

17. Long ZW, Yu HM, Wang YN, Liu D, Chen YZ, Zhao YX, et al. Association of IL-17 polymorphisms with gastric cancer risk in Asian populations. World journal of gastroenterology. 2015;21(18):5707-18.

18. Yang S, Lee J, Choi JI, Kim YW, Ryu KW, Sung J, et al. Effects of alcohol consumption, ALDH2 rs671 polymorphism, and Helicobacter pylori infection on the gastric cancer risk in a Korean population. Oncotarget. 2017;8(4):6630-41.

19. Park B, Yang S, Lee J, Woo HD, Choi JI, Kim YW, et al. Genome-Wide Association of Genetic Variation in the PSCA Gene with Gastric Cancer Susceptibility in a Korean Population. Cancer research and treatment : official journal of Korean Cancer Association. 2019;51(2):748-57.

20. Tanikawa C, Kamatani Y, Toyoshima O, Sakamoto H, Ito H, Takahashi A, et al. Genome-wide association study identifies gastric cancer susceptibility loci at 12q24.11-12 and 20q11.21. 2018;109(12):4015-24.

21. 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(9):764-7.
22. Sakamoto H, Yoshimura K, Saeki N, Katai H, Shimoda T, Matsuno Y, et al. Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. Nature genetics. 2008;40(6):730-40.

23. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. The New England journal of medicine. 2000;343(2):78-85.

24. Park SL, Cheng I, Haiman CA. Genome-Wide Association Studies of Cancer in Diverse Populations. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2018;27(4):405-17.

25. Nagini S. Carcinoma of the stomach: A review of epidemiology, pathogenesis, molecular genetics and chemoprevention. World journal of gastrointestinal oncology. 2012;4(7):156-69.

26. Woodcock CC, Huang Y, Woodcock SR, Salvatore SR, Singh B, Golin-Bisello F, et al. Nitro-fatty acid inhibition of triple-negative breast cancer cell viability, migration, invasion, and tumor growth. The Journal of biological chemistry. 2018;293(4):1120-37.

27. Wu Y, Zhang J, Zheng Y, Ma C, Liu XE, Sun X. miR-216a-3p Inhibits the Proliferation, Migration, and Invasion of Human Gastric Cancer Cells via Targeting RUNX1 and Activating the NF-κB Signaling Pathway. Oncology research. 2018;26(1):157-71.

28. Liu X, Zhang P, Bao Y, Han Y, Wang Y, Zhang Q, et al. Zinc finger protein ZBTB20 promotes Toll-like receptor-triggered innate immune responses by repressing IκBα gene transcription. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(27):11097-102.

29. Dong LM, Potter JD, White E, Ulrich CM, Cardon LR, Peters U. Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. Jama. 2008;299(20):2423-36.

30. Yusefi AR, Bagheri Lankarani K, Bastani P, Radinmanesh M, Kavosi Z. Risk Factors for Gastric Cancer: A Systematic Review. Asian Pacific journal of cancer prevention : APJCP. 2018;19(3):591-603.

31. Mocellin S, Verdi D, Pooley KA, Nitti D. Genetic variation and gastric cancer risk: a field synopsis and meta-analysis. Gut. 2015;64(8):1209-19.

32. He Y, Wang C, Wang Z, Zhou Z. Genetic variant PLCE1 rs2274223 and gastric cancer: more to be explored? Gut. 2016;65(2):359-60.
33. Sung H, Yang HH, Hu N, Su H, Taylor PR, Hyland PL. Functional annotation of high-quality SNP biomarkers of gastric cancer susceptibility: the Yin Yang of PSCA rs2294008. Gut. 2016;65(2):361-4.

34. Negovan A, Iancu M, Fülöp E, Bănescu C. Helicobacter pylori and cytokine gene variants as predictors of premalignant gastric lesions. World journal of gastroenterology. 2019;25(30):4105-24.

35. Yoon H, Kim N. Diagnosis and management of high risk group for gastric cancer. Gut and liver. 2015;9(1):5-17.

36. Ladeiras-Lopes R, Pereira AK, Nogueira A, Pinheiro-Torres T, Pinto I, Santos-Pereira R, et al. Smoking and gastric cancer: systematic review and meta-analysis of cohort studies. Cancer causes & control : CCC. 2008;19(7):689-701.

37. La Torre G, Chiaradia G, Gianfagna F, De Lauretis A, Boccia S, Mannocci A, et al. Smoking status and gastric cancer risk: an updated meta-analysis of case-control studies published in the past ten years. Tumori. 2009;95(1):13-22.

38. Tramacere I, Negri E, Pelucchi C, Bagnardi V, Rota M, Scotti L, et al. A meta-analysis on alcohol drinking and gastric cancer risk. Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23(1):28-36.

39. Lope V, Fernández de Larrea N, Pérez-Gómez B, Martín V, Moreno V, Costas L, et al. Menstrual and Reproductive Factors and Risk of Gastric and Colorectal Cancer in Spain. PloS one. 2016;11(10):e0164620.

40. 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. International journal of cancer. 2015;136(5):E359-86.

41. Haemmerle M, Stone RL, Menter DG, Afshar-Kharghan V, Sood AK. The Platelet Lifeline to Cancer: Challenges and Opportunities. Cancer cell. 2018;33(6):965-83.

42. Chen L, Hao Y, Cong X, Zou M, Li S, Zhu L, et al. Peripheral Venous Blood Platelet-to-Lymphocyte Ratio (PLR) for Predicting the Survival of Patients With Gastric Cancer Treated With SOX or XELOX Regimen Neoadjuvant Chemotherapy. Technology in cancer research & treatment. 2019;18:1533033819829485.

43. Ye DM, Xu G, Ma W, Li Y, Luo W, Xiao Y, et al. Significant function and research progress of biomarkers in gastric cancer. Oncology letters. 2020;19(1):17-29.
**Supplemental table 1** The primers for ZBTB20 gene amplification and sequencing.

| SNP      | 1st-PCR primer | 2nd-PCR primer | UEP-DIR | UEP-SEQ                      |
|----------|----------------|----------------|---------|------------------------------|
| rs10934270 | ACGTTGGATGTCATCACCTGATGTGTC | ACGTTGGATGGAATTAAACCAGATCGGGCGG | F       | GCCATTGTATATATTTCACA         |
| rs9288999 | ACGTTGGATGATGGTCTGAGAGGATTAAAAAG | ACGTTGGATGATCGTCTATTGGGCTGCTTTC | R       | atgTAGTGAATGTGCTACACAC       |
| rs9841504 | ACGTTGGATGGAAATCTGTTAAGGAGC | ACGTTGGATGACATTTCAGGCCCCTATACCTTC | F       | TCCCTTTTTTCTATATATTTTTTA    |
| rs73230612 | ACGTTGGATGCTAGGGCAATGACATGTTGT | ACGTTGGATGGTAAACGAAAGGCAAGCA | R       | CAAGCAAATATGAAAAAGCAATGA    |

SNP, single nucleotide polymorphism;
UEP-DIR, Unique extension primer-direction;
UEP-SEQ, Unique extension primer-sequencing.
### Supplemental table 2 The correlation between SNPs and gastric cancer sensitivity in subgroups (pathological grade and lymph node metastasis).

| SNP ID   | Model    | Genotype | Pathological grade | Lymph node metastasis |
|----------|----------|----------|--------------------|------------------------|
|          |          |          | Case III & IV | Control I & II | OR (95% CI) | p       | Case Yes | Control No | OR (95% CI) | p       |
| rs10934270 | Allele   | T        | 52      | 21           | 1.15 (0.67-1.95) | 0.619   | 49       | 19         | 1.07 (0.61-1.87) | 0.807    |
|          |          | C        | 426     | 197          | 1.00             |         | 421      | 175        | 1.00             |         |
|          | Genotype | TT       | 4       | 2            | 1.01 (0.18-5.67) | 0.993   | 3        | 1          | 1.32 (0.13-13.01) | 0.810    |
|          |          | TC       | 44      | 17           | 1.31 (0.70-2.44) | 0.403   | 43       | 17         | 1.13 (0.60-2.12) | 0.699    |
|          |          | CC       | 191     | 90           | 1.00             |         | 189      | 79         | 1.00             |         |
|          | Dominant | TT-TC    | 48      | 19           | 1.27 (0.70-2.32) | 0.428   | 46       | 18         | 1.14 (0.62-2.11) | 0.668    |
|          |          | CC       | 191     | 90           | 1.00             |         | 189      | 79         | 1.00             |         |
|          | Recessive| TT       | 4       | 2            | 0.96 (0.17-5.40) | 0.966   | 3        | 1          | 1.30 (0.13-12.70) | 0.824    |
|          |          | TC-CC    | 235     | 107          | 1.00             |         | 232      | 96         | 1.00             |         |
|          | Log-additive | -     | -       | -            | 1.20 (0.71-2.04) | 0.495   | -        | -          | 1.14 (0.65-1.99) | 0.655    |
| rs9288999 | Allele   | G        | 170     | 78           | 1.00 (0.71-1.39) | 0.987   | 165      | 65         | 1.08 (0.76-1.54) | 0.667    |
|          |          | A        | 306     | 140          | 1.00             |         | 303      | 129        | 1.00             |         |
|          | Genotype | GG       | 26      | 16           | 0.77 (0.37-1.59) | 0.478   | 27       | 10         | 1.13 (0.50-2.54) | 0.778    |
|          |          | GA       | 118     | 46           | 1.27 (0.77-2.07) | 0.349   | 111      | 45         | 1.07 (0.64-1.76) | 0.807    |
|          |          | AA       | 94      | 47           | 1.00             |         | 96       | 42         | 1.00             |         |
|          | Dominant | GG-GA    | 144     | 62           | 1.14 (0.72-1.81) | 0.588   | 138      | 55         | 1.08 (0.66-1.74) | 0.766    |
|          |          | AA       | 94      | 47           | 1.00             |         | 96       | 42         | 1.00             |         |
|          | Recessive| GG       | 26      | 16           | 0.68 (0.35-1.34) | 0.265   | 27       | 10         | 1.09 (0.50-2.36) | 0.831    |
|          |          | GA-AA    | 212     | 93           | 1.00             |         | 207      | 87         | 1.00             |         |
|          | Log-additive | -     | -       | -            | 0.98 (0.69-1.37) | 0.885   | -        | -          | 1.06 (0.74-1.53) | 0.746    |
| rs9841504 | Allele   | G        | 66      | 38           | 0.76 (0.49-1.17) | 0.214   | 67       | 35         | 0.76 (0.48-1.18) | 0.219    |
|          |          | C        | 412     | 180          | 1.00             |         | 403      | 159        | 1.00             |         |
| Genotype   | Allele | Count | p<0.05  | Allele | Count | p<0.05  | Allele | Count | p<0.05  |
|------------|--------|-------|---------|--------|-------|---------|--------|-------|---------|
| GG         | C      | 210   | 1.00(0.73-1.39) | 96     | 0.984 | 0.37(0.10-1.33) | 5     | 0.64 | 0.128   |
| GC         | T      | 266   | 1.00   | 122    | 0.94  | 0.64(0.43-1.26) | 28    | 0.266| 0.452   |
| CC         |        |       |        |        |       |         |        |       |         |
| Dominant   | GG-GC  | 58    | 0.72(0.43-1.20) | 33     | 0.209 | 0.88(0.50-1.52) | 28    | 0.266| 0.379   |
| CC         |        | 181   | 1.00   | 76     | 1.00  |         |        |       |         |
| Recessive  | GG     | 8     | 0.69(0.22-2.20) | 55     | 0.530 | 0.38(0.11-1.37) | 5     | 0.37 | 0.128   |
| GC         | CC     | 231   | 1.00   | 104    | 1.00  |         |        |       |         |
| Log-additive | -     | -     | 0.76(0.50-1.16) | -     | 0.208 | 0.76(0.49-1.17) | -     | -     | 0.212   |
| rs73230612 | Allele |        |        |        |       |         |        |       |         |
| C          |        | 210   | 1.00(0.73-1.39) | 96     | 0.984 | 0.37(0.10-1.33) | 5     | 0.64 | 0.128   |
| T          |        | 266   | 1.00   | 122    | 0.94  | 0.64(0.43-1.26) | 28    | 0.266| 0.452   |
| Genotype   | CC     | 45    | 1.04(0.53-2.01) | 21     | 0.919 | 0.84(0.43-1.66) | 43    | 0.84 | 0.621   |
| CT         |        | 120   | 1.08(0.64-1.82) | 54     | 0.775 | 1.14(0.65-1.98) | 45    | 1.14 | 0.647   |
| TT         |        | 73    | 1.00   | 34     | 1.00  |         |        |       |         |
| Dominant   | CC-CT  | 165   | 1.07(0.65-1.75) | 75     | 0.798 | 1.04(0.62-1.75) | 67    | 0.879| 1.00    |
| TT         |        | 73    | 1.00   | 34     | 1.00  |         |        |       |         |
| Recessive  | CC     | 45    | 0.99(0.55-1.77) | 21     | 0.965 | 0.78(0.44-1.40) | 43    | 0.78 | 0.401   |
| CT-TT      |        | 193   | 1.00   | 88     | 1.00  |         |        |       |         |
| Log-additive | -     | -     | 1.03(0.74-1.43) | -     | 0.885 | 0.94(0.67-1.32) | -     | -     | 0.704   |

SNP: Single nucleotide polymorphisms;
OR, Odds ratio;
CI, Confidence interval;
p < 0.05 indicates statistical significance;
"-" indicates Log-additive model.
### Supplemental table 3 Clinical characteristics of patients based on the genotypes of selected SNPs.

| Characteristics | rs10934270 | rs9288999 |
|-----------------|------------|-----------|
|                 | TT         | TC        | CC        | p  | AA         | AG         | GG         | p  |
| CEA             | 16.83 ± 9.26 | 15.8 ± 8.07 | 17.41 ± 11.07 | 0.582 | 17.56 ± 11.83 | 16.9 ± 9.31 | 16.66 ± 12.06 | 0.846 |
| TNF (fmol/ml)   | 0.87 ± 0.07 | 0.89 ± 0.08 | 1.04 ± 2.3 | 0.877 | 1.19 ± 3.29 | 0.89 ± 0.06 | 0.9 ± 0.09 | 0.459 |
| CA50 (U/ml)     | 13.52 ± 15.32 | 8.29 ± 13.96 | 7.2 ± 11.5 | 0.385 | 8.47 ± 13.86 | 7.3 ± 10.98 | 4.5 ± 8.38 | 0.251 |
| CA19-9 (U/ml)   | 75.59 ± 117.45 | 53.91 ± 105.74 | 45.01 ± 89.03 | 0.607 | 52.15 ± 145.46 | 47.87 ± 90.62 | 24.71 ± 21.62 | 0.326 |
| CA242 (KU/ml)   | 12.06 ± 16.71 | 16.33 ± 28 | 15.53 ± 30.41 | 0.949 | 13.52 ± 27.15 | 17.83 ± 32.05 | 12 ± 26.79 | 0.0385 |
| WBC (L)         | 5.74 ± 2.72 | 6.27 ± 4.06 | 6.73 ± 6.52 | 0.905 | 7.64 ± 8.04 | 6.18 ± 4.32 | 5 ± 2.76 | 0.139 |
| HGB (g/L)       | 118.33 ± 29.67 | 109.46 ± 27.85 | 103.36 ± 23.78 | 0.304 | 104.93 ± 26.32 | 105.18 ± 24.56 | 101.92 ± 19.9 | 0.845 |
| PLT (L)         | 166 ± 110.31 | 211.25 ± 89.15 | 200.38 ± 110.29 | 0.807 | 213.55 ± 118.84 | 202.94 ± 103.18 | 161.24 ± 65.01 | 0.149 |

CEA, carcinoembryonic antigen;
TNF, tumor necrosis factor;
CA50, carbohydrate antigen 50;
CA19-9, carbohydrate antigen 19-9;
CA242, carbohydrate antigen 242;
WBC, white blood cells;
HGB, hemoglobin;
PLT, platelet.
Table title:

**Table 1** Characteristics of patients with gastric cancer (GC) and healthy individuals.

**Table 2** The basic information and HWE about the selected SNPs of ZBTB20.

**Table 3** Analysis of the association between susceptibility of gastric cancer and single nucleotide polymorphism of ZBTB20.

**Table 4** The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (age and gender).

**Table 5** The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (smoking and drinking).

**Table 6** The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (BMI).

**Table 7** The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (adenocarcinoma).

**Table 8** SNP–SNP interaction models analyzed by the MDR method.

**Table 9** Clinical characteristics of patients based on the genotypes of selected SNPs.
The titles and legends of figures:

**Title: Figure 1** Dendrogram analysis of SNP-SNP interaction.

**Figure legends** The colors in the tree diagram represent synergy (yellow) or redundancy (blue).

**Title: Figure 2** Multifactor dimensionality reduction (MDR) analysis for SNPs (10934270, rs9288999, rs9841504 and rs73230612) of *ZBTB20* interaction.

**Figure legends** In each box, the left bar represents cases and the right bar represents controls. The light gray boxes indicate the low risk of gastric cancer and dark gray boxes indicate the high risk, the empty boxes mean no data.
Table 1 Characteristics of patients with gastric cancer (GC) and healthy individuals.

| Characteristics                  | Cases | Control | p     |
|----------------------------------|-------|---------|-------|
|                                 | n = 509 | n = 507 |       |
| Age (years)                      | Mean ± SD | 61.12 ± 11.33 | 61.35 ± 8.84 | 0.712 |
|                                 | > 60 | 279 (55%) | 315 (62%)  |       |
|                                 | ≤ 60 | 230 (45%) | 192 (38%)  |       |
| Gender                           | Male | 382 (75%) | 379 (75%)  | 0.942 |
|                                 | Female | 127 (25%) | 128 (25%)  |       |
| Lymph node metastasis            | Yes | 235 (46%) | -         |       |
|                                 | No | 97 (19%) | -         |       |
| pathological grade               | III & IV | 239 (47%) | -         |       |
|                                 | I & II | 109 (21%) | -         |       |
| Adenocarcinoma                   | - | 314 (62%) | -         |       |
| Smoking                          | Yes | 233 (56%) | 114 (22%)  |       |
|                                 | No | 270 (53%) | 172 (34%)  |       |
| Drinking                         | Yes | 133 (26%) | 119 (23%)  |       |
|                                 | No | 357 (70%) | 142 (28%)  |       |
| BMI (kg/m²)                      | BMI > 24 | 72 (14%) | 183 (36%)  |       |
|                                 | BMI ≤ 24 | 401 (79%) | 170 (34%)  |       |

BMI: body mass index.
| Gene     | SNP ID       | Role  | Chr: Position | Alleles (A/B) | MAF Cases | MAF Controls | HWE (p Value) | Haploreg 4.1                                                                 |
|----------|--------------|-------|---------------|---------------|-----------|--------------|---------------|-----------------------------------------------------------------------------|
| ZBTB20   | rs10934270   | Intron| 3: 114384900  | T/C           | 0.102     | 0.095        | 0.434         | SiPhy cons; DNase; Motifs changed; Selected eQTL hits                       |
| ZBTB20   | rs9288999    | Intron| 3: 114429080  | G/A           | 0.357     | 0.432        | 0.928         | Enhancer histone marks; Motifs changed                                      |
| ZBTB20   | rs9841504    | Intron| 3: 114643917  | G/C           | 0.145     | 0.134        | 0.848         | Enhancer histone marks; Motifs changed; NHGRI/EBI GWAS hits                 |
| ZBTB20   | rs73230612   | Intron| 3: 115131989  | C/T           | 0.442     | 0.416        | 0.201         | Motifs changed                                                              |

HWE: Hardy–Weinberg equilibrium;

$p > 0.05$ indicates that the genotypes were in Hard-Weinberg Equilibrium;

SNP: Single nucleotide polymorphisms;

MAF: minor allele frequency.
Table 3 Analysis of the association between susceptibility of gastric cancer and single nucleotide polymorphism of \textit{ZBTB20}.

| SNP ID | Model | Genotype | Case | Control | Adjusted by age and gender | OR (95% CI) | p   |
|--------|-------|----------|------|---------|----------------------------|-------------|-----|
|        |       |          |      |         | OR (95% CI) | 0.571 |      |
| rs10934270 | Allele | T      | 104  | 96      | 1.09 (0.81 - 1.46) | 0.777 |      |
|         | C       | 914    | 918  | 1.00    |              |      |      |
|         | Genotype | TT   | 7    | 6       | 1.17 (0.39 - 3.52) | 0.629 |      |
|         |          | TC   | 90   | 84      | 1.08 (0.78 - 1.50) | 0.595 |      |
|         |          | CC   | 412  | 417     | 1.00           |      |      |
|         | Dominant | TT-TC | 97   | 90      | 1.09 (0.79 - 1.50) | 0.595 |      |
|         |          | CC   | 412  | 417     | 1.00           |      |      |
|         | Recessive | TT  | 7    | 6       | 1.16 (0.39 - 3.47) | 0.796 |      |
|         |          | TC-CC | 502  | 501     | 1.00           |      |      |
|         | Log-additive | -    | -    | -       | 1.08 (0.81 - 1.45) | 0.583 |      |
| rs9288999 | Allele | G      | 363  | 438     | 0.73 (0.61 - 0.87) | 0.001* |      |
|         | A       | 653    | 576  | 1.00    |              |      |      |
|         | Genotype | GG   | 57   | 95      | 0.48 (0.33 - 0.71) | 0.0003* |      |
|         |          | GA   | 249  | 248     | 0.82 (0.62 - 1.07) | 0.140 |      |
|         |          | AA   | 202  | 164     | 1.00           |      |      |
|         | Dominant | GG-GA | 306  | 343     | 0.72 (0.56 - 0.94) | 0.014* |      |
|         |          | AA   | 202  | 164     | 1.00           |      |      |
|         | Recessive | GG  | 57   | 95      | 0.55 (0.38 - 0.78) | 0.001* |      |
|         |          | GA-AA | 451  | 412     | 1.00           |      |      |
|         | Log-additive | -    | -    | -       | 0.72 (0.60 - 0.87) | 0.0005* |      |
| rs9841504 | Allele | G      | 148  | 136     | 1.10 (0.85 - 1.41) | 0.464 |      |
|         | C       | 870    | 878  | 1.00    |              |      |      |
|         | Genotype | GG   | 14   | 8       | 1.76 (0.73 - 4.25) | 0.208 |      |
|         |          | GC   | 120  | 120     | 1.01 (0.76 - 1.35) | 0.935 |      |
|         |          | CC   | 375  | 379     | 1.00           |      |      |
|         | Dominant | GG-GC | 134  | 128     | 1.06 (0.80 - 1.40) | 0.688 |      |
|         |          | CC   | 375  | 379     | 1.00           |      |      |
|         | Recessive | GG  | 14   | 8       | 1.76 (0.73 - 4.23) | 0.209 |      |
|         |          | GC-CC | 495  | 499     | 1.00           |      |      |
|         | Log-additive | -    | -    | -       | 1.10 (0.86 - 1.41) | 0.466 |      |
| rs73230612 | Allele | C      | 449  | 422     | 1.11 (0.93 - 1.32) | 0.241 |      |
|         | T       | 567    | 592  | 1.00    |              |      |      |
|         | Genotype | CC   | 98   | 95      | 1.19 (0.83 - 1.69) | 0.349 |      |
|         |          | CT   | 253  | 232     | 1.25 (0.94 - 1.65) | 0.119 |      |
|         |          | TT   | 157  | 180     | 1.00           |      |      |
|         | Dominant | CC-CT | 351  | 327     | 1.23 (0.95 - 1.60) | 0.121 |      |
|         |          | TT   | 157  | 180     | 1.00           |      |      |
|         | Recessive | CC  | 98   | 95      | 1.04 (0.76 - 1.42) | 0.809 |      |
|         |          | CT-TT | 410  | 412     | 1.00           |      |      |
|         | Log-additive | -    | -    | -       | 1.11 (0.93 - 1.32) | 0.244 |      |
SNP: Single nucleotide polymorphisms;
OR, Odds ratio;
CI, Confidence interval;
$p < 0.05$ indicates statistical significance.
Table 4: The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (age and gender).

| SNP ID   | Model | genotype | Age, years | Gender | OR (95% CI) | p     | OR (95% CI) | p     |
|----------|-------|----------|------------|--------|-------------|-------|-------------|-------|
|          |       |          | OR (95% CI) < 60 | p     | OR (95% CI) ≤ 60 | p     | OR (95% CI) Female | p     | OR (95% CI) Male | p     |
| rs10934270 | Allele | T        | 1.24(0.84-1.81) | 0.280 | 0.91(0.58-1.43) | 0.681 | 1.01(0.55-1.85) | 0.978 | 1.11(0.80-1.55) | 0.530 |
|          |       | C        | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Genotype | TT      | 0.82(0.15-4.58) | 0.820 | 2.10(0.40-11.11) | 0.384 | 1.01(0.06-16.46) | 0.993 | 1.20(0.36-3.98) | 0.767 |
|          |       | TC      | 1.43(0.03-2.21) | 0.106 | 0.75(0.44-1.25) | 0.269 | 1.01(0.52-1.96) | 0.979 | 1.11(0.76-1.62) | 0.586 |
|          |       | CC      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Dominant | TT-TT   | 1.39(0.91-2.12) | 0.130 | 0.82(0.50-1.34) | 0.422 | 1.01(0.53-1.93) | 0.978 | 1.12(0.78-1.61) | 0.552 |
|          |       | CC      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Recessive | TT      | 0.77(0.14-4.28) | 0.763 | 2.02(0.42-11.63) | 0.353 | 1.01(0.06-16.40) | 0.994 | 1.18(0.36-3.90) | 0.789 |
|          |       | CC      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Log-additive | G     | 1.30(0.88-1.93) | 0.184 | 0.91(0.59-1.41) | 0.667 | 1.01(0.55-1.85) | 0.978 | 1.11(0.80-1.53) | 0.545 |
|          |       | A       | 0.83(0.65-1.05) | 0.116 | 0.59(0.45-0.78) |       | 0.82(0.57-1.17) | 0.271 | 0.70(0.57-0.87) |       |
|          | Genotype | GG      | 0.65(0.38-1.12) | 0.122 | 0.33(0.18-0.59) |       | 0.61(0.29-1.30) | 0.199 | 0.45(0.28-0.70) |       |
|          |       | GA      | 0.83(0.58-1.18) | 0.298 | 0.75(0.48-1.17) |       | 0.96(0.56-1.65) | 0.881 | 0.77(0.57-1.06) | 0.108 |
|          |       | AA      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Dominant | GG-GA   | 0.79(0.56-1.11) | 0.171 | 0.60(0.40-0.91) |       | 0.86(0.51-1.44) | 0.567 | 0.68(0.51-0.92) | 0.108 |
|          |       | AA      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Recessive | GG      | 0.73(0.44-1.20) | 0.213 | 0.39(0.23-0.65) | 0.003* | 0.62(0.31-1.24) | 0.179 | 0.52(0.34-0.78) | 0.002* |
|          |       | GA-AA   | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Log-additive | G     | 0.81(0.63-1.04) | 0.105 | 0.60(0.45-0.79) |       | 0.82(0.57-1.17) | 0.266 | 0.69(0.56-0.86) |       |
|          |       | C       | 1.04(0.75-1.43) | 0.820 | 1.26(0.83-1.90) | 0.276 | 1.22(0.74-2.00) | 0.429 | 1.06(0.79-1.42) | 0.701 |
|          | Genotype | GG      | 1.16(0.38-3.57) | 0.797 | 6.47(0.79-52.99) | 0.082 | 2.62(0.49-13.93) | 0.259 | 1.50(0.53-4.26) | 0.450 |
|          |       | GC      | 1.11(0.76-1.62) | 0.587 | 1.00(0.62-1.61) | 0.993 | 1.04(0.58-1.87) | 0.890 | 1.00(0.72-1.40) | 0.998 |
|          |       | CC      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Dominant | GG-GC   | 1.12(0.77-1.61) | 0.563 | 1.13(0.71-1.80) | 0.602 | 1.14(0.65-2.00) | 0.640 | 1.03(0.74-1.43) | 0.852 |
|          |       | CC      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Recessive | GG      | 1.13(0.37-3.45) | 0.835 | 6.48(0.79-52.89) | 0.081 | 2.60(0.49-13.74) | 0.262 | 1.50(0.53-4.25) | 0.449 |
|          |       | GC-GC   | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | Log-additive | G     | 1.10(0.79-1.53) | 0.566 | 1.25(0.83-1.88) | 0.294 | 1.21(0.75-1.96) | 0.441 | 1.06(0.79-1.42) | 0.701 |
|          |       | C       | 1.12(0.89-1.41) | 0.331 | 1.09(0.83-1.43) | 0.542 | 0.86(0.61-1.23) | 0.412 | 1.21(0.99-1.48) | 0.068 |
|          | Genotype | CC      | 1.09(0.68-1.75) | 0.714 | 1.17(0.66-2.05) | 0.597 | 0.75(0.37-1.53) | 0.435 | 1.38(0.92-2.08) | 0.123 |
|          |       | CT      | 1.28(0.88-1.86) | 0.190 | 1.19(0.77-1.85) | 0.436 | 0.86(0.50-1.50) | 0.599 | 1.42(1.03-1.96) | 0.035 |
|          |       | TT      | 1.00        |        | 1.00         |       | 1.00         |       | 1.00         |       |
|          | CC-CT | TT   | CC-CT | TT   | CC-CT | TT   | CC-CT | TT   | CC-CT | TT   |
|----------|-------|------|-------|------|-------|------|-------|------|-------|------|
| Dominant | 1.22  | 0.256| 1.18  | 0.427| 0.83  | 0.481| 1.41  | 0.586|
| Recessive| 0.95  | 0.801| 1.05  | 0.858| 0.82  | 0.543| 1.13  | 0.526|
| Log-additive | 1.00 | 1.00 | 1.00 | 1.00 | 0.87 | 0.422| 1.21  | 0.070|

SNP: Single nucleotide polymorphisms; OR, Odds ratio; CI, Confidence interval; \( p < 0.05 \) indicates statistical significance; “-” indicates Log-additive model.
### Table 5: The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (smoking and drinking).

| SNP ID      | Model | genotype | Smoking |          | OR (95% CI) | p  | OR (95% CI) | p  |
|-------------|-------|----------|---------|----------|-------------|-----|-------------|-----|
|            |       |          | Yes     | OR No    |             |     |             |     |
| rs10934270 | Allele| T        | 0.90(0.56-1.46) | 0.666 | 0.86(0.54-1.36) | 0.523 | 0.91(0.54-1.56) | 0.744 | 0.97(0.61-1.55) | 0.897 |
|            |       | C        | 1.00    |          |             |     |             |     |
| Genotype   |       | TT       | 1.30(0.26-6.59) | 0.754 | 0.29(0.03-3.25) | 0.315 | 2.04(0.39-10.85) | 0.401 | 0.42(0.06-3.07) | 0.395 |
|            |       | TC       | 0.76(0.43-1.33) | 0.335 | 0.93(0.56-1.53) | 0.763 | 0.72(0.37-1.38) | 0.317 | 1.08(0.64-1.83) | 0.775 |
|            |       | CC       | 1.00    |          |             |     |             |     |
| Dominant   |       | TT-TT    | 0.80(0.46-1.38) | 0.419 | 0.89(0.54-1.45) | 0.629 | 0.82(0.44-1.51) | 0.525 | 1.03(0.62-1.72) | 0.916 |
|            |       | CC-CC    | 1.00    |          |             |     |             |     |
| Recessive  |       | TT       | 1.37(0.27-6.92) | 0.705 | 0.29(0.03-3.29) | 0.320 | 2.17(0.41-11.48) | 0.361 | 0.42(0.06-3.03) | 0.388 |
|            |       | TC-CC    | 1.00    |          |             |     |             |     |
| Log-additive |       |          |          |          |             |     |             |     |
| rs9288999  | Allele| T        | 0.87(0.55-1.39) | 0.565 | 0.85(0.54-1.36) | 0.505 | 0.95(0.57-1.58) | 0.829 | 0.98(0.61-1.57) | 0.923 |
|            |       | C        | 0.78(0.56-1.08) | 0.132 | 0.77(0.58-1.01) | 0.059 | 0.75(0.52-1.08) | 0.120 | 0.79(0.59-1.04) | 0.905 |
| Genotype   |       | GG       | 0.54(0.27-1.07) | 0.077 | 0.46(0.26-0.82) | 0.009* | 0.49(0.23-1.07) | 0.072 | 0.54(0.30-0.96) | 0.035* |
|            |       | GA       | 0.84(0.51-1.39) | 0.498 | 1.16(0.76-1.78) | 0.500 | 1.04(0.60-1.8) | 0.904 | 1.02(0.66-1.58) | 0.925 |
|            |       | AA       | 1.00    |          |             |     |             |     |
| Dominant   |       | GG-GA    | 0.75(0.47-1.21) | 0.237 | 0.92(0.62-1.38) | 0.690 | 0.86(0.51-1.43) | 0.555 | 0.87(0.58-1.31) | 0.513 |
|            |       | AA-CC    | 1.00    |          |             |     |             |     |
| Recessive  |       | GG       | 0.59(0.32-1.11) | 0.101 | 0.42(0.25-0.72) | 0.002* | 0.48(0.24-0.99) | 0.047 | 0.53(0.31-0.90) | 0.018* |
|            |       | GA-AA    | 1.00    |          |             |     |             |     |
| Log-additive |       |          |          |          |             |     |             |     |
| rs9841504  | Allele| T        | 0.76(0.54-1.05) | 0.095 | 0.76(0.57-1.00) | 0.052 | 0.77(0.53-1.10) | 0.150 | 0.78(0.59-1.04) | 0.093 |
|            |       | C        | 1.49(0.88-2.51) | 0.137 | 1.02(0.70-1.47) | 0.928 | 1.20(0.69-2.06) | 0.520 | 1.02(0.69-1.50) | 0.937 |
| Genotype   |       | GG       | /        | 0.999 | 2.00(0.53-7.46) | 0.305 | /            | 0.999 | 1.48(0.41-5.39) | 0.551 |
|            |       | GC       | 1.43(0.80-2.56) | 0.234 | 0.87(0.56-1.35) | 0.542 | 1.06(0.58-1.95) | 0.840 | 0.92(0.59-1.46) | 0.732 |
|            |       | CC       | 1.00    |          |             |     |             |     |
| Dominant   |       | GG-GC    | 1.53(0.86-2.73) | 0.150 | 0.94(0.62-1.44) | 0.778 | 1.14(0.62-2.06) | 0.678 | 0.97(0.62-1.50) | 0.883 |
|            |       | GC-CC    | 1.00    |          |             |     |             |     |
| Recessive  |       | GG       | /        | 0.999 | 2.07(0.56-7.69) | 0.279 | /            | 0.999 | 1.51(0.42-5.47) | 0.530 |
|            |       | GC-CC    | 1.00    |          |             |     |             |     |
| Log-additive |       |          |          |          |             |     |             |     |
| rs73230612 | Allele| T        | 1.59(0.92-2.76) | 0.100 | 1.02(0.71-1.47) | 0.914 | 1.21(0.69-2.14) | 0.512 | 1.02(0.70-1.48) | 0.937 |
|            |       | C        | 1.09(0.80-1.5) | 0.581 | 1.15(0.87-1.52) | 0.326 | 1.37(0.96-1.95) | 0.080 | 1.12(0.85-1.49) | 0.417 |
| Genotype   |       | CC       | 1.18(0.62-2.33) | 0.618 | 1.21(0.67-2.17) | 0.524 | 2.03(0.97-4.25) | 0.060 | 1.17(0.66-2.05) | 0.595 |
|            |       | CT       | 1.16(0.68-1.98) | 0.590 | 1.29(0.85-1.97) | 0.233 | 1.52(0.83-2.77) | 0.175 | 1.38(0.90-2.13) | 0.142 |
|            |       | TT       | 1.00    |          |             |     |             |     |
|         | CC-CT  | TT     | CC   | CT-TT | Log-additive |
|---------|--------|--------|------|-------|--------------|
| Dominant| 1.16(0.70-1.92) | 0.552  | 1.27(0.85-1.90) | 0.237 | 1.65(0.93-2.93) | 0.085 | 1.32(0.88-1.97) | 0.177 |
| Recessive| 1.07(0.62-1.86) | 0.799  | 1.05(0.61-1.79) | 0.868 | 1.54(0.83-2.88) | 0.172 | 0.98(0.58-1.64) | 0.932 |
| Log-additive| 1.09(0.79-1.50) | 0.600  | 1.14(0.86-1.51) | 0.362 | 1.43(0.99-2.07) | 0.057 | 1.13(0.86-1.50) | 0.382 |

OR, Odds ratio; CI, Confidence interval; 
\( p < 0.05 \) indicates statistical significance. 
“-” indicates Log-additive model; 
“/” indicates data missing.
Table 6 The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (BMI).

| SNP ID  | Model   | genotype | BMI < 24 | OR (95% CI) | p   | BMI > 24 | OR (95% CI) | p   |
|---------|---------|----------|----------|-------------|-----|----------|-------------|-----|
| rs10934270 | Allele  | T/C      | 0.87(0.58-1.30) | 0.490 | 1.26(0.66-2.41) | 0.490 |
|          |         |          | 0.78(0.19-3.19) | 0.726 | 2.90(0.18-47.36) | 0.456 |
|          |         |-network  | 0.83(0.52-1.32) | 0.425 | 1.17(0.56-2.42) | 0.680 |
|          |         |         | 0.82(0.52-1.29) | 0.396 | 1.22(0.60-2.48) | 0.584 |
|          |         |         | 0.80(0.20-3.30) | 0.762 | 2.82(0.17-45.89) | 0.467 |
|          | Log-additive | -      | 0.84(0.56-1.26) | 0.401 | 1.26(0.65-2.43) | 0.500 |
| rs9288999 | Allele  | G/A      | 0.68(0.53-0.89) | 0.004* | 0.80(0.54-1.19) | 0.265 |
|          |         |          | 0.41(0.24-0.71) | 0.001* | 0.57(0.24-1.39) | 0.220 |
|          |         |          | 0.84(0.56-1.26) | 0.389 | 0.77(0.42-1.41) | 0.391 |
|          |         |          | 0.70(0.48-1.23) | 0.372 | 0.72(0.40-1.29) | 0.269 |
|          |         |          | 0.46(0.28-0.74) | 0.002* | 0.68(0.30-1.52) | 0.342 |
|          | Log-additive | -      | 0.67(0.51-0.88) | 0.004* | 0.76(0.50-1.15) | 0.198 |
| rs9841504 | Allele  | G/C      | 1.41(0.95-2.10) | 0.0087 | 1.43(0.95-2.10) | 0.0180 |
|          |         |          | 3.56(0.44-29.1) | 0.236 | 11.19(1.31-108.82) | 0.028* |
|          |         |          | 1.32(0.84-2.05) | 0.227 | 0.87(0.45-1.69) | 0.688 |
|          |         |          | 1.38(0.89-2.14) | 0.149 | 1.11(0.60-2.04) | 0.744 |
|          |         |          | 3.35(0.41-27.33) | 0.259 | 12.29(1.36-111.21) | 0.026* |
|          | Log-additive | -      | 1.40(0.93-2.10) | 0.105 | 1.35(0.80-2.28) | 0.267 |
| rs73230612 | Allele  | C/T      | 1.16(0.90-1.50) | 0.256 | 0.98(0.67-1.45) | 0.937 |
|          |         |          | 1.37(0.80-2.34) | 0.248 | 0.89(0.38-2.06) | 0.777 |
|          |         |          | 1.13(0.75-1.70) | 0.558 | 1.39(0.74-2.61) | 0.302 |
|          |         |          | 1.19(0.81-1.75) | 0.374 | 1.24(0.68-2.26) | 0.482 |
|          |         |          | 1.27(0.79-2.04) | 0.319 | 0.72(0.34-1.51) | 0.385 |
|          | Log-additive | -      | 1.16(0.90-1.51) | 0.251 | 1.00(0.67-1.48) | 0.987 |

SNP: Single nucleotide polymorphisms; OR, Odds ratio; CI, Confidence interval; p < 0.05 indicates statistical significance; "-" indicates Log-additive model.
Table 7  The SNPs of ZBTB20 associated with susceptibility of gastric cancer in the subgroup tests (adenocarcinoma).

| SNP ID   | Model       | genotype  | Adenocarcinoma (patients with GC) | OR (95% CI) | p     |
|----------|-------------|-----------|-----------------------------------|-------------|-------|
|          |             | Case      | Control                          |             |       |
|          |             | Yes       | No                               |             |       |
| rs10934270 | Allele      | T/C       | 60 96                            | 1.01(0.72-1.42) | 0.954 |
|          | Homozygote  | TT/CC     | 5 6                             | 1.30(0.39-4.32) | 0.667 |
|          | Heterozygote| TC        | 50 84                          | 0.95(0.65-1.40) | 0.804 |
|          | Dominant    | TT-TC/CC  | 55 90                          | 0.98(0.67-1.41) | 0.898 |
|          | Recessive   | TT/TC-CC  | 5 6                            | 1.31(0.40-4.35) | 0.656 |
|          | Log-additive| -         | -                               | 1.00(0.72-1.39) | 0.996 |
| rs9288999  | Allele      | G/A       | 222 438                        | 0.72(0.59-0.89) | 0.002*|
|          | Homozygote  | GG/AA     | 40 95                          | 0.52(0.34-0.81) | 0.004*|
|          | Heterozygote| GA        | 142 248                        | 0.72(0.53-0.98) | 0.037*|
|          | Dominant    | GG-GA/AA  | 182 343                        | 0.67(0.50-0.89) | 0.006*|
|          | Recessive   | GG/GA-AA  | 40 95                          | 0.63(0.42-0.94) | 0.024*|
|          | Log-additive| -         | -                               | 0.72(0.59-0.89) | 0.002*|
| rs9841504  | Allele      | G/C       | 93 136                          | 1.12(0.84-1.49) | 0.427 |
|          | Homozygote  | GG/CC     | 8 8                             | 1.68(0.62-4.53) | 0.309 |
|          | Heterozygote| GC        | 77 120                          | 1.07(0.77-1.49) | 0.697 |
|          | Dominant    | GG-GC/CC  | 85 128                          | 1.11(0.80-1.52) | 0.538 |
|          | Recessive   | GG/GC-CC  | 8 8                             | 1.65(0.61-4.44) | 0.322 |
|          | Log-additive| -         | -                               | 1.13(0.85-1.50) | 0.405 |
| rs73230612 | Allele      | C/T       | 282 422                        | 1.15(0.94-1.41) | 0.173 |
|          | Homozygote  | CC/TT     | 64 95                           | 1.28(0.86-1.92) | 0.224 |
|          | Heterozygote| CT        | 154 232                        | 1.26(0.91-1.73) | 0.164 |
|          | Dominant    | CC-CT/TT  | 218 327                        | 1.26(0.93-1.71) | 0.128 |
|          | Recessive   | CC/CT-TT  | 64 95                           | 1.12(0.79-1.60) | 0.523 |
|          | Log-additive| -         | -                               | 1.15(0.94-1.40) | 0.174 |

SNP: Single nucleotide polymorphisms; OR, Odds ratio; CI, Confidence interval; p < 0.05 indicates statistical significance; "-" indicates Log-additive model.

Downloaded from http://portlandpress.com/bioscirep/article-pdf/doi/10.1042/BSR20202102/893253/bsr-2020-2102.pdf by guest on 20 September 2020

Bioscience Reports. This is an Accepted Manuscript. You are encouraged to use the Version of Record, which will replace this version. The most up-to-date version is available at https://doi.org/10.1042/BSR20202102
### Table 8 SNP–SNP interaction models analyzed by the MDR method.

| Model                                      | Training Bal. Acc | Testing Bal. Acc | OR (95% CI)          | p value | CVC |
|--------------------------------------------|-------------------|------------------|----------------------|---------|-----|
| rs9288999                                  | 0.540             | 0.521            | 1.86 (1.30-2.65)     | 0.0006  | 10/10 |
| rs9288999, rs73230612                      | 0.557             | 0.535            | 1.60 (1.24-2.05)     | 0.0003  | 10/10 |
| rs10934270, rs9288999, rs73230612           | 0.569             | 0.500            | 1.72 (1.33-2.22)     | < 0.0001| 7/10  |
| rs10934270, rs9288999, rs9841504, rs73230612| 0.586             | 0.517            | 1.96 (1.52-2.51)     | < 0.0001| 10/10 |

MDR, multifactor dimensionality reduction; Bal. Acc., balanced accuracy; CVC, cross-validation consistency; OR, odds ratio; 95% CI, 95% confidence interval. 

p values were calculated using χ² tests; p < 0.05 indicates statistical significance.
Table 9 Clinical characteristics of patients based on the genotypes of selected SNPs.

| Characteristics | rs9841504 |     |     |     | rs73230612 |     |     |     |
|-----------------|----------|-----|-----|-----|------------|-----|-----|-----|
|                 | CC       | CG  | GG  | p   | TT         | TC  | CC  | p   |
| CEA             | 17.28 ± 10.59 | 16.99 ± 11.16 | 14.06 ± 3.49 | 0.637 | 17.38 ± 8.85 | 16.24 ± 9.34 | 18.94 ± 15.26 | 0.230 |
| TNF (fmol/ml)   | 0.90 ± 0.07 | 1.39 ± 4.30 | 0.88 ± 0.06 | 0.196 | 0.88 ± 0.07 | 1.12 ± 2.92 | 0.90 ± 0.07 | 0.569 |
| CA50 (U/ml)     | 7.48 ± 11.94 | 7.74 ± 12.76 | 6.05 ± 7.78 | 0.916 | 8.35 ± 13.58 | 7.02 ± 11.04 | 7.45 ± 12.01 | 0.681 |
| CA19-9 (U/ml)   | 45.57 ± 89.00 | 55.45 ± 107.75 | 20.19 ± 8.30 | 0.491 | 51.65 ± 91.31 | 44.68 ± 94.12 | 44.63 ± 90.63 | 0.827 |
| CA242 (KU/ml)   | 16.48 ± 30.97 | 14.11 ± 27.37 | 5.47 ± 7.72 | 0.493 | 23.51 ± 42.37 | 11.39 ± 19.71 | 12.87 ± 23.18 | 0.005* |
| WBC (L)         | 6.27 ± 5.92 | 7.83 ± 6.67 | 3.83 ± 2.13 | 0.293 | 8.14 ± 8.17 | 5.96 ± 5.45 | 6.16 ± 3.66 | 0.127 |
| HGB (g/L)       | 102.76 ± 22.62 | 111.02 ± 28.04 | 93.5 ± 37.03 | 0.108 | 105.38 ± 26.76 | 102.33 ± 22.16 | 108.7 ± 26.79 | 0.417 |
| PLT (L)         | 189.64 ± 94.02 | 229.40 ± 132.05 | 307.00 ± 89.10 | 0.048* | 203.16 ± 99.62 | 207.62 ± 117.15 | 187.68 ± 94.63 | 0.669 |

CEA, carcinoembryonic antigen; TNF, tumor necrosis factor; CA50, carbohydrate antigen 50; CA19-9, carbohydrate antigen 19-9; CA242, carbohydrate antigen 242; WBC, white blood cells; HGB, hemoglobin; PLT, platelet.
### Supplemental table 1 The primers for ZBTB20 gene amplification and sequencing.

| SNP         | 1st-PCR primer                  | 2nd-PCR primer                  | UEP-DIR | UEP-SEQ                  |
|-------------|--------------------------------|--------------------------------|---------|--------------------------|
| rs10934270  | ACGTTGGATGGCCATCCTAGTATGTGTC   | ACGTTGGATGGAAATTTAACCAGATCGGC   | F       | GCATTTGATGTTATATTTTCACA  |
| rs9288999   | ACGTTGGATATGTTGCTGGTAGTAAGAAG  | ACGTTGGATGTGAAATTAACCAGATCGGG   | R       | atgtTAGTTAATGCTACAACAC   |
| rs9841504   | ACGTTGGATGGAATCTCACTGTTAAGGAGC | ACGTTGGATGACATTTTGGCCGTACCTTC   | F       | TCCCTTTTTCCCTATATATTTTA  |
| rs73230612  | ACGTTGGATGCTAGGCCATGACCATGTGT  | ACGTTGGATGGTAAACGAAAGGCAAGCA    | R       | CAAGCAATATGAAAAAGCAATGA  |

SNP: single nucleotide polymorphism;

UEP-DIR: Unique extension primer-direction;

UEP-SEQ: Unique extension primer-sequencing.
Supplemental table 2 The correlation between SNPs and gastric cancer sensitivity in subgroups (pathological grade and lymph node metastasis).

| SNP ID    | Model         | Genotype | Pathological grade | Lymph node metastasis |
|-----------|---------------|----------|--------------------|-----------------------|
|           |               |          | Case   | Control | OR (95% CI) | p      | Case | Control | OR (95% CI) | p      |
| rs10934270| Allele        | T        | III & IV | 52      | 21       | 1.15(0.67-1.95) | 0.619 | 49   | 19       | 1.07(0.61-1.87) | 0.807 |
|           |               | C        | I & II   | 426     | 197      | 1.00            |       | 421  | 175      | 1.00            |       |
|           | Genotype      |          |          |         |          |                 |       |      |          |                 |       |
|           |               | TT       |          | 4       | 2        | 1.01(0.18-5.67) | 0.993 | 3    | 1        | 1.32(0.13-13.01) | 0.810 |
|           |               | TC       |          | 44      | 17       | 1.31(0.70-2.44) | 0.403 | 43   | 17       | 1.13(0.60-2.12)  | 0.699 |
|           |               | CC       |          | 191     | 90       | 1.00            |       | 189  | 79       | 1.00            |       |
|           | Dominant      |          |          |         |          |                 |       |      |          |                 |       |
|           |               | TT-TC    |          | 48      | 19       | 1.27(0.70-2.32) | 0.428 | 46   | 18       | 1.14(0.62-2.11)  | 0.668 |
|           |               | CC       |          | 191     | 90       | 1.00            |       | 189  | 79       | 1.00            |       |
|           | Recessive     |          |          |         |          |                 |       |      |          |                 |       |
|           |               | TT       |          | 4       | 2        | 0.96(0.17-5.40) | 0.966 | 3    | 1        | 1.30(0.13-12.70) | 0.824 |
|           |               | TC-CC    |          | 235     | 107      | 1.00            |       | 232  | 96       | 1.00            |       |
|           | Log-additive  |          |          |         |          |                 |       |      |          |                 |       |
|           |               |          |          |         |          | 1.20(0.71-2.04) | 0.495 | -    | -        | 1.14(0.65-1.99)  | 0.655 |
| rs9288999 | Allele        | G        |          | 170     | 78       | 1.00(0.71-1.39) | 0.987 | 165  | 65       | 1.08(0.76-1.54)  | 0.667 |
|           |               | A        |          | 306     | 140      | 1.00            |       | 303  | 129      | 1.00            |       |
|           | Genotype      |          |          |         |          |                 |       |      |          |                 |       |
|           |               | GG       |          | 26      | 16       | 0.77(0.37-1.59) | 0.478 | 27   | 10       | 1.13(0.50-2.54)  | 0.778 |
|           |               | GA       |          | 118     | 46       | 1.27(0.77-2.07) | 0.349 | 111  | 45       | 1.07(0.64-1.76)  | 0.807 |
|           |               | AA       |          | 94      | 47       | 1.00            |       | 96   | 42       | 1.00            |       |
|           | Dominant      |          |          |         |          |                 |       |      |          |                 |       |
|           |               | GG-GA    |          | 144     | 62       | 1.14(0.72-1.81) | 0.588 | 138  | 55       | 1.08(0.66-1.74)  | 0.766 |
|           |               | AA       |          | 94      | 47       | 1.00            |       | 96   | 42       | 1.00            |       |
|           | Recessive     |          |          |         |          |                 |       |      |          |                 |       |
|           |               | GG       |          | 26      | 16       | 0.68(0.35-1.34) | 0.265 | 27   | 10       | 1.09(0.50-2.36)  | 0.831 |
|           |               | GA-AA    |          | 212     | 93       | 1.00            |       | 207  | 87       | 1.00            |       |
|           | Log-additive  |          |          |         |          |                 |       |      |          |                 |       |
|           |               |          |          |         |          | 0.98(0.69-1.37) | 0.885 | -    | -        | 1.06(0.74-1.53)  | 0.746 |
| rs9841504 | Allele        | G        |          | 66      | 38       | 0.76(0.49-1.17)| 0.214 | 67   | 35       | 0.76(0.48-1.18)  | 0.219 |
|           |               | C        |          | 412     | 180      | 1.00            |       | 403  | 159      | 1.00            |       |
| Genotype    | Allele | Sample Size | Frequency   | OR (95% CI) |
|-------------|--------|-------------|-------------|-------------|
| GG          | C      | 8           | 0.64 (0.20-2.05) | 0.452       | 0.37 (0.10-1.33) | 0.128   |
| GG          | T      | 50          | 0.74 (0.43-1.26) | 0.266       | 0.88 (0.50-1.52) | 0.641   |
| CC          | C      | 181         | 1.00         | 0.37 (0.10-1.33) | 0.128   |
| CC          | T      | 76          | 1.00         | 0.38 (0.11-1.37) | 0.139   |

| Dominant    |        |             |             |             |
|-------------|--------|-------------|-------------|-------------|
| GG-GC       | C      | 58          | 0.72 (0.43-1.20) | 0.209       | 0.79 (0.47-1.33) | 0.379   |
| GG-GC       | T      | 33          | 1.00         | 0.76 (0.49-1.17) | 0.212   |
| CC          | C      | 181         | 1.00         | 0.93 (0.66-1.30) | 0.661   |
| CC          | T      | 76          | 1.00         | 0.90 (0.63-1.30) | 0.661   |

| Recessive   |        |             |             |             |
|-------------|--------|-------------|-------------|-------------|
| GG          | C      | 8           | 0.69 (0.22-2.20) | 0.530       | 0.38 (0.11-1.37) | 0.139   |
| GG          | T      | 5           | 0.76 (0.50-1.16) | 0.208       | 0.76 (0.49-1.17) | 0.212   |
| GC          | C      | 231         | 1.00         | 0.96 (0.66-1.30) | 0.661   |
| GC          | T      | 104         | 1.00         | 0.94 (0.67-1.32) | 0.704   |

| Log-additive|        |             |             |             |
|-------------|--------|-------------|-------------|-------------|
| rs73230612  | C      | 210         | 1.00 (0.73-1.39) | 0.984       | 0.93 (0.66-1.30) | 0.661   |
| rs73230612  | T      | 96          | 1.00         | 0.93 (0.66-1.30) | 0.661   |

| Genotype    | Allele | Sample Size | Frequency   | OR (95% CI) |
|-------------|--------|-------------|-------------|-------------|
| CC          | C      | 45          | 1.04 (0.53-2.01) | 0.919       | 0.84 (0.43-1.66) | 0.621   |
| CC          | T      | 21          | 1.08 (0.64-1.82) | 0.775       | 1.14 (0.65-1.98) | 0.647   |
| TT          | C      | 73          | 1.00         | 0.71 (0.43-1.21) | 0.465   |
| TT          | T      | 34          | 1.00         | 0.71 (0.43-1.21) | 0.465   |

| Dominant    |        |             |             |             |
|-------------|--------|-------------|-------------|-------------|
| CC-CT       | C      | 165         | 1.07 (0.65-1.75) | 0.798       | 1.04 (0.62-1.75) | 0.879   |
| CC-CT       | T      | 75          | 1.00         | 0.71 (0.43-1.21) | 0.465   |

| Recessive   |        |             |             |             |
|-------------|--------|-------------|-------------|-------------|
| CC          | C      | 45          | 0.99 (0.55-1.77) | 0.965       | 0.78 (0.44-1.40) | 0.401   |
| CC          | T      | 21          | 1.00         | 0.94 (0.67-1.32) | 0.704   |
| CT-TT       | C      | 193         | 1.00         | 0.94 (0.67-1.32) | 0.704   |
| CT-TT       | T      | 88          | 1.00         | 0.94 (0.67-1.32) | 0.704   |
## Supplemental table 3 Clinical characteristics of patients based on the genotypes of selected SNPs.

| Characteristics | rs10934270 | rs9288999 |
|-----------------|------------|-----------|
|                 | TT         | TC        | CC         | p     | AA         | AG         | GG         | p     |
| CEA             | 16.83 ± 9.26 | 15.8 ± 8.07 | 17.41 ± 11.07 | 0.582 | 17.56 ± 11.83 | 16.9 ± 9.31 | 16.66 ± 12.06 | 0.846 |
| TNF (fmol/ml)   | 0.87 ± 0.07 | 0.89 ± 0.08 | 1.04 ± 2.3 | 0.877 | 1.19 ± 3.29 | 0.89 ± 0.06 | 0.9 ± 0.09 | 0.459 |
| CA50 (U/ml)     | 13.52 ± 15.32 | 8.29 ± 13.96 | 7.2 ± 11.5 | 0.385 | 8.47 ± 13.86 | 7.3 ± 10.98 | 4.5 ± 8.38 | 0.251 |
| CA19-9 (U/ml)   | 75.59 ± 117.45 | 53.91 ± 105.74 | 45.01 ± 89.03 | 0.607 | 52.15 ± 105.46 | 47.87 ± 90.62 | 24.71 ± 21.62 | 0.326 |
| CA242 (KU/ml)   | 12.06 ± 16.71 | 16.33 ± 28 | 15.53 ± 30.41 | 0.949 | 13.52 ± 27.15 | 17.83 ± 32.05 | 12 ± 26.79 | 0.00385 |
| WBC (L)         | 5.74 ± 2.72 | 6.27 ± 4.06 | 6.73 ± 6.52 | 0.905 | 7.64 ± 8.04 | 6.18 ± 4.32 | 5 ± 2.76 | 0.139 |
| HGB (g/L)       | 118.33 ± 29.67 | 109.46 ± 27.85 | 103.36 ± 23.78 | 0.304 | 104.93 ± 26.32 | 105.18 ± 24.56 | 101.92 ± 19.9 | 0.845 |
| PLT (L)         | 166 ± 110.31 | 211.25 ± 89.15 | 200.38 ± 110.29 | 0.807 | 213.55 ± 118.84 | 202.94 ± 103.18 | 161.24 ± 65.01 | 0.149 |

CEA, carcinoembryonic antigen;
TNF, tumor necrosis factor;
CA50, carbohydrate antigen 50;
CA19-9, carbohydrate antigen 19-9;
CA242, carbohydrate antigen 242;
WBC, white blood cells;
HGB, hemoglobin;
PLT, platelet.