The Impact of BCL11A Polymorphisms on Endometrial Cancer Risk Among Chinese Han Females

Junhong Cai, Siyuan Peng, Haibo Wang, Shan Bao

1Medical Laboratory Center, Hainan General Hospital/Hainan Affiliated Hospital of Hainan Medical University, Haikou, Hainan, 570311, People’s Republic of China; 2Department of Gynaecology and Obstetrics, Hainan General Hospital/Hainan Affiliated Hospital of Hainan Medical University, Haikou, Hainan, 570311, People’s Republic of China

Correspondence: Shan Bao, Tel/Fax +86-0898-68642629, Email baoshan3@126.com

Background: Endometrial carcinoma (EC) is one of the most common malignant gynecological malignancies. BCL11A gene may have a tumor-suppressor role in EC. Until now, no studies have reported the effect of BCL11A variants on EC predisposition in Chinese population.

Methods: Six BCL11A polymorphisms were genotyped using Agena MassARRAY system among 509 EC patients and 506 matched healthy women. Risk assessment of the BCL11A polymorphisms for EC risk was performed by calculating odds ratios (OR) with 95% confidence intervals (CI) through logistic regression models.

Results: We found that rs7581162 (OR = 1.29, p = 0.012), rs10189857 (OR = 1.26, p = 0.028), rs1427407 (OR = 1.30, p = 0.015), rs766432 (OR = 1.27, p = 0.025), and rs6729815 (OR = 1.32, p = 0.008) in BCL11A were associated with higher susceptibility to EC in Chinese Han women. Age and BMI stratified analysis displayed that the risk association between BCL11A variants and EC predisposition might be age- and BMI-dependent. Haplotype analysis revealed that A rs10189857 T rs1427407 and G rs10189857 G rs1427407 haplotypes were related to an increased risk of EC. MDR analysis indicated that rs1427407 was the most influential attributor on EC risk in the single-locus model, and the best combination was the two-locus model containing rs7581162 and rs766432.

Conclusion: Our study provided the first evidence that rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 in BCL11A were risk factors for EC in Chinese Han women. These findings add our understanding of the role of BCL11A gene in EC pathogenesis.

Keywords: endometrial carcinoma, BCL11A variants, susceptibility, haplotype analysis, MDR analysis

Introduction

Endometrial carcinoma (EC) occurs in the endometrium characterized by a thickening or mass of the endometrium. EC is one of the most common malignant gynecological tumors with an estimated age-standardized incidence risk of 13.1. In China, EC ranks fourth among female cancers in both incidence and mortality, and the incidence increases steadily with decreasing age at diagnosis. Risk factors for EC are unopposed estrogen, family history of EC, age, excess body weight, diabetes, hypertension, and Lynch syndrome. Although several risk factors have been identified, endometrial carcinogenesis remains poorly understood. Genetic variants such as single nucleotide polymorphisms (SNPs) are known to play important roles in cancer predisposition. Many SNPs related to EC susceptibility have been identified in previous genome-wide association studies (GWAS), but many polymorphic loci have not been reported.

The B-cell lymphoma/leukemia 11A (BCL11A) gene encodes a C2H2 type zinc finger protein transcription factor, and the role of BCL11A in tumors appears to be contextual. In some cancers, it has oncogenic effects, while in some cancers, it may act as a tumor suppressor. For example, BCL11A, as an oncogene, promoted tumor formation, cancer cell mobility and epithelial-mesenchymal transition by activating the Wnt/β-catenin signaling pathway in breast cancer carcinogenesis. Downregulation of BCL11A protein in colorectal cancer cells was related to enhanced radioresistance, supporting BCL11A as a tumor-suppressor role. The expression of BCL11A in ER−/PR−EC was higher than that in
normal endometrium, atypical hyperplasia endometrium, and ER+/PR+EC. Moreover, the expression of BCL11A in EC was associated with age, menopause, EC classification, para-aortic lymph node metastasis, tumor differentiation, histological type, ER/PR expression and p53 expression. A targeted next-generation sequencing study displayed that BCL11A mutations were associated with endometriosis and tumor lesions. In EC, consistent with an anti-cancer function, credible variants in BCL11A were associated with reduced BCL11A expression in endometrial tumors. A GWAS study in an Australian population found that BCL11A polymorphism was associated with EC risk. Until now, there were no studies reported the effect of BCL11A genetic variants on EC predisposition in Chinese population.

Here, six SNPs (rs7581162, rs10189857, rs1427407, rs766432, rs6729815, and rs2556378) in BCL11A with a minor allele frequency (MAF) > 0.05 in dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP) and the call rate >95% in our study population were randomly selected and genotyped to evaluate the potential association between BCL11A variants and EC risk among the Chinese Han women at single-SNP or combined SNPs interface.

**Materials and Methods**

**Subjects**

A total of 509 patients with EC and 506 matched healthy women were recruited from Hainan General Hospital (Table 1). All participants were genetically unrelated Han Chinese women. Patients were diagnosed with EC based on histopathologic biopsies using the guidelines of the International Federation of Obstetrics and Gynecology (FIGO) criteria. Patients who received preoperative chemotherapy, radiotherapy, or hormone therapy were excluded. Patients with immunological diseases or other cancers were also excluded. Age-matched healthy control underwent routine gynecologic examination at the same hospital. Selection criteria included no malignancy, no history of cancer and no chronic or acute disease. Demographical and clinical characteristics were collected from medical records. This study was approved by the ethics committee of Hainan General Hospital (Ethical approval No.: Med-Eth-Re [2018] 14), in compliance with the Declaration of Helsinki. Informed consent was obtained from all recruited participants.

**SNP Selection and Genotyping**

Six SNPs (rs7581162, rs10189857, rs1427407, rs766432, rs6729815, and rs2556378) in the BCL11A gene with minor allele frequency (MAF) > 0.05 in the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP) and call rate >95% were selected. The potential functions of these polymorphisms were evaluated through the HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php), RegulomeDB (https://regulome.stanford.edu/regulome-search/) and GTEx Portal databases (https://gtexportal.org/home/).

| Characteristics | Cases (n = 509) | Control (n = 506) | p-value |
|-----------------|----------------|-----------------|---------|
| Age (years, mean ± SD) | 54.94 ± 8.85 | 54.61 ± 9.07 | 0.553 |
| BMI (kg/m²) | | | |
| ≥ 24 | 199 (39.1%) | 191 (37.7%) | 0.699 |
| < 24 | 310 (60.9%) | 315 (62.3%) | |
| CEA (ng/mL, mean ± SD) | 11.20 ± 2.09 | 2.08 ± 2.22 | < 0.001 |
| AFP (ng/mL, mean ± SD) | 8.96 ± 5.32 | 2.86 ± 1.17 | < 0.001 |
| CA199 (U/mL, mean ± SD) | 17.87 ± 20.99 | 12.79 ± 10.45 | < 0.001 |
| CA125 (U/mL, mean ± SD) | 24.22 ± 39.00 | 12.97 ± 9.51 | < 0.001 |
| Stage | | | |
| I–II | 261 (51.3%) | | |
| III–IV | 93 (18.3%) | | |
| Missing | 155 (30.5%) | | |

**Notes:** p values were calculated using χ² tests or Student’s t-test. Bold indicated that p < 0.05 meant the data was statistically significant.

**Abbreviations:** SD, standard deviation; BMI, body mass index; CEA, carcinoembryonic antigen; AFP, alpha fetoprotein; CA, carbohydrate antigen.
Venous blood (5 mL) was collected from each subject in an EDTA vacutainer tube. Genomic DNA was isolated using a commercially available GoldMag DNA Purification Kit (GoldMag Co. Ltd, Xi’an, China). Genotyping was performed by the Agena MassARRAY system (Agena, San Diego, CA, USA) as previously described. Primers design (Suppl. Table 1) and data interpretation were performed by the corresponding supporting software following the manufacturer’s instructions. The MassARRAY platform is based on MALDI-TOF (matrix-assisted laser desorption/ionization—time of flight) mass spectrometry. The analytical accuracy of MALDI-TOF MS is quite high, 0.1–0.01% of the determined mass. Furthermore, 5% of the samples were used for re-genotyping to quality control, and the consistency rate was 100%.

**Statistical Analysis**

Age and clinical characteristics were expressed as mean ± standard deviation (SD), and differences between EC patients and health controls were evaluated by Student’s t-test. The difference of body mass index (BMI) between cases and controls were analyzed by χ²-test Hardy–Weinberg equilibrium (HWE) was assessed to compare the genotype frequencies in controls using the goodness-of-fit chi-square test. Comparison of genotype and allele frequencies in cases and controls were performed by χ²-test. The major-type allele was used as a reference, and minor-type allele was used as a risk factor. Risk assessment of BCL11A polymorphisms for EC risk was performed by calculating odds ratios (OR) with 95% confidence intervals (CI) using logistic regression models. D’ values for pairwise linkage disequilibrium (LD) plots were generated by Haploview software (version 4.2). The relationship of BCL11A haplotypes with EC susceptibility was evaluated by χ²-test and logistic regression model. Multifactor dimensionality reduction (MDR) analysis was performed using MDR_3.0.2 software to identify high-order interaction models for EC predisposition. False-positive report probability (FPRP) analysis was used to evaluate the noteworthy associations and statistical power of the significant findings. We set 0.2 as a FPRP threshold and assigned a priori probability of 0.1 for an association with genotypes under investigation. Analysis of Variance (ANOVA) was performed for the correlation of BCL11A variants with CEA, AFP, CA199, CA125 of EC patients and healthy controls. Data analysis was performed using SPSS version 20.0 (SPSS Inc., Chicago, IL, USA) and PLINK version 1.0.7 software. p-value was two-tailed and p < 0.05 was defined statistical significance, whereas adjusted p < 0.05/5 was considered significant after Bonferroni correction.

**Results**

**Subject Characteristics**

In the study, we enrolled 509 EC patients (54.94 ± 8.85 years) and 506 healthy controls (54.61 ± 9.07 years), as shown in Table 1. Age and BMI distributions were not significantly different between cases and controls (p = 0.553, and p = 0.669, respectively). There were significant difference in the levels of CEA, AFP, CA199, and CA125 between two groups (p < 0.001). There were 261 cases of I–II stage and 93 cases of III–IV stage.

**Relationships of BCL11A Polymorphisms with EC Predisposition**

Six SNPs in BCL11A were genotyped for subsequent studies, and the call rates were > 99.5%, as summarized in Table 2. The genotype distribution of rs2556378 polymorphism was not inconsistent with HWE (p = 0.001), therefore, this variant was excluded from subsequent studies. The MAFs of rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 in EC patients were higher than those in healthy controls. BCL11A rs7581162-T (OR = 1.29, 95% CI: 1.06–1.57, p = 0.012), rs10189857-A (OR = 1.26, 95% CI: 1.03–1.54, p = 0.028), rs1427407-T (OR = 1.30, 95% CI: 1.05–1.60, p = 0.015), rs766432-C (OR = 1.27, 95% CI: 1.03–1.56, p = 0.025), and rs6729815-T (OR = 1.32, 95% CI: 1.07–1.62, p = 0.008) was associated with the increased risk of EC. The significance of rs6729815 still existed after Bonferroni correction. The potential function of these polymorphisms by HaploReg v4.1 database and RegulomeDB database was displayed in Table 2. Based on GTEx Portal database, the genotypes of rs1427407 (p = 0.00026) was related to the mRNA expression of BCL11A in cell (Suppl. Figure 1).

Multiple genetic models were used to evaluate the relationship of BCL11A polymorphisms to EC predisposition. After adjusting for age, rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 were associated with higher EC susceptibility (Table 3). Specifically, the TT genotype of rs7581162 (genotype: TT vs AA, OR = 2.16, 95% CI: 1.27–3.69, p = 0.005;
Table 2 Basic Characteristics and Allele Model About Candidate SNPs in the BCL11A Gene

| SNPs ID    | Chr: Position | Allele (Minor/Major) | Call Rate | O(HET) | E(HET) | $p^a$-value for HWE | MAF   | Case OR (95% CI) | Control OR (95% CI) | $p^b$ | Allele Model                                                                 |
|------------|---------------|----------------------|-----------|--------|--------|-------------------|-------|-----------------|---------------------|------|-----------------------------------------------------------------------------|
| rs7581162  | 2:60,477,349  | T/A                  | 99.9%     | 0.384  | 0.362  | 0.219             | 0.287 | 0.238           | 1.29 (1.06–1.57)     | 0.012| Promoter histone marks, Enhancer histone marks, Motifs changed, GRASP QTL hits |
| rs10189857 | 2:60,486,100  | A/G                  | 99.8%     | 0.345  | 0.346  | 1.000             | 0.264 | 0.222           | 1.26 (1.03–1.54)     | 0.028| DNAse, Motifs changed, NHGRI/EBI GWAS hits                                  |
| rs1427407  | 2:60,490,908  | T/G                  | 99.8%     | 0.310  | 0.321  | 0.486             | 0.246 | 0.200           | 1.30 (1.05–1.60)     | 0.015| Enhancer histone marks, Proteins bound, Motifs changed, NHGRI/EBI GWAS hits, GRASP QTL hits |
| rs766432   | 2:60,492,835  | C/A                  | 100%      | 0.338  | 0.335  | 0.895             | 0.254 | 0.212           | 1.27 (1.03–1.56)     | 0.025| Enhancer histone marks, DNAse, Motifs changed, NHGRI/EBI GWAS hits, GRASP QTL hits |
| rs6729815  | 2:60,496,537  | T/C                  | 100%      | 0.330  | 0.335  | 0.790             | 0.262 | 0.212           | 1.32 (1.07–1.62)     | 0.008*| Enhancer histone marks, Motifs Changed, NHGRI/EBI GWAS hits, Promoter histone marks, Enhancer histone marks, DNAse, NHGRI/EBI GWAS hits, GRASP QTL hits |
| rs2556378  | 2:60,535,367  | T/G                  | 100%      | 0.405  | 0.352  | 0.001             | /     | /               | /                   | /    | eQTL + TF binding + DNase peak                                             |

Notes: $p^a$-values for the HWE test were calculated using goodness of fit $\chi^2$ test. $p^b$ were calculated using by Fisher’s exact test. Bold indicated that $p < 0.05$ meant the data was statistically significant. *p indicate that after Bonferroni correction ($p < 0.05/5$) means the data is statistically significant.

Abbreviations: SNP, single-nucleotide polymorphism; MAF, minor allele frequency; O(HET), observed heterozygosity; E(HET), expected heterozygosity; HWE, Hardy–Weinberg equilibrium.
recessive: TT vs AA-AT, OR = 2.03, 95% CI: 1.20–3.42, \( p = 0.008 \); and additive: AA+AT+TT, OR = 1.30, 95% CI: 1.07–1.59, \( p = 0.010 \) and AA genotype of rs10189857 (genotype: AA vs GG, OR = 1.78, 95% CI: 1.05–3.01, \( p = 0.033 \); and additive: GG+GA+AA, OR = 1.26, 95% CI: 1.03–1.54, \( p = 0.027 \)) contributed to increased EC risk. Rs1427407 had a risk-effect for EC under the genotype (GT vs GG: OR = 1.33, 95% CI: 1.02–1.74, \( p = 0.033 \), dominant (GT-TT vs GG: OR = 1.37, 95% CI: 1.06–1.76, \( p = 0.016 \)) and additive (GG+GT+TT: OR = 1.30, 95% CI: 1.05–1.60, \( p = 0.014 \)) models. Rs766432 was associated with increased EC predisposition (dominant: AC-CC vs AA, OR = 1.32, 95% CI: 1.03–1.70, \( p = 0.029 \); and additive: AA+AC+CC, OR = 1.28, 95% CI: 1.04–1.57, \( p = 0.022 \)). Moreover, rs6729815 might have higher risk for the occurrence EC in the genotype (TT vs CC: OR = 1.94, 95% CI: 1.14–3.29, \( p = 0.015 \), dominant (CT-TT vs CC: OR = 1.32, 95% CI: 1.03–1.70, \( p = 0.030 \)), recessive (TT vs CC-CT: OR = 1.79, 95% CI: 1.06–3.02, \( p = 0.028 \)) and additive (TT+CC+CT: OR = 1.31, 95% CI: 1.07–1.61, \( p = 0.009 \)) models. The significance of rs7581162 in the genotype and recessive models and rs6729815 in the additive model still existed after Bonferroni correction.

### Stratified Analysis for the Contribution of BCL11A Variant to EC Risk

We further explored the stratified analysis by age and BMI for the relationships of BCL11A variants with EC risk (Table 4). In the subjects with age > 55 years/BMI ≥ 24 kg/m\(^2\), no significant association between BCL11A polymorphism with the susceptibility to EC occurrence was observed. Among those aged 55 or younger and those had BMI < 24 kg/m\(^2\), we found that rs7581162-T, rs10189857-A, rs1427407-T, rs766432-C, and rs6729815-T were risk factors for the development of EC.

The results of age stratification (age ≤ 55 years) were as follows: in the allele (rs7581162, OR = 1.40, \( p = 0.010 \); rs10189857, OR = 1.38, \( p = 0.016 \); rs1427407, OR = 1.44, \( p = 0.008 \); rs766432, OR = 1.42, \( p = 0.010 \); and rs6729815, OR = 1.48, \( p = 0.004 \), respectively), genotype (rs7581162, OR = 2.84, \( p = 0.001 \); rs10189857, OR = 2.53, \( p = 0.005 \); rs1427407, OR = 2.37, \( p = 0.012 \); rs766432, OR = 2.55, \( p = 0.008 \); and rs6729815, OR = 2.70, \( p = 0.003 \), respectively), dominant (rs1427407, OR = 1.41, \( p = 0.043 \)), recessive (rs7581162, OR = 2.78, \( p = 0.001 \); rs10189857, OR = 2.44, \( p = 0.006 \); rs1427407, OR = 2.18, \( p = 0.022 \); rs766432, OR = 2.35, \( p = 0.014 \); rs6729815, OR = 2.52, \( p = 0.004 \), respectively) and additive (rs7581162, OR = 1.38, \( p = 0.012 \); rs10189857, OR = 1.36, \( p = 0.020 \); rs1427407, OR = 1.41, \( p = 0.011 \); rs766432, OR = 1.40, \( p = 0.012 \); rs6729815, OR = 1.43, \( p = 0.006 \), respectively) models. The significance of rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 among those aged 55 or younger still existed after Bonferroni correction.

The associated results of BMI stratification (BMI < 24 kg/m\(^2\)) were as follows: rs7581162 in the allele (OR = 1.42, \( p = 0.006 \)), genotype (OR = 2.41, \( p = 0.009 \)), dominant (OR = 1.45, \( p = 0.022 \)), recessive (OR = 2.14, \( p = 0.022 \)), and additive (OR = 1.44, \( p = 0.005 \)) models; rs10189857 in the allele (OR = 1.37, \( p = 0.017 \)), genotype (OR = 1.43, \( p = 0.037 \)), dominant (OR = 1.47, \( p = 0.018 \)), and additive (OR = 1.38, \( p = 0.016 \)) models; rs1427407 in the allele (OR = 1.41, \( p = 0.012 \)), genotype (OR = 1.57, \( p = 0.009 \)), dominant (OR = 1.57, \( p = 0.007 \)), and additive (OR = 1.41, \( p = 0.012 \)) models; rs766432 in the allele (OR = 1.37, \( p = 0.019 \)), genotype (OR = 1.52, \( p = 0.014 \)), dominant (OR = 1.53, \( p = 0.010 \)), and additive (OR = 1.39, \( p = 0.015 \)) models; rs6729815 in the allele (OR = 1.50, \( p = 0.003 \)), genotype (OR = 1.57, \( p = 0.009 \), and OR = 1.99, \( p = 0.042 \)), dominant (OR = 1.63, \( p = 0.003 \)), and additive (OR = 1.49, \( p = 0.003 \)) models. The significance of rs7581162, rs1427407, and rs6729815 among those with BMI < 24 kg/m\(^2\) still existed after Bonferroni correction.

The association of these SNPs with the stage of EC patients was assessed (Suppl. Table 3). However, no significant association was detected between BCL11A SNPs and at age in EC patients.

### The Association Between BCL11A Haplotypes and EC Susceptibility

Moreover, haplotype analysis was performed to estimate the association between BCL11A haplotypes and EC susceptibility. As shown in Figure 1, rs10189857 and rs1427407 are in linkage disequilibrium. Haplotype analysis revealed that A\(_{rs10189857}\) T\(_{rs1427407}\) (adjusted OR = 1.30, 95% CI: 1.05–1.60, \( p = 0.016 \)) and G\(_{rs10189857}\) G\(_{rs1427407}\) (adjusted OR = 1.26, 95% CI: 1.03–1.54, \( p = 0.023 \)) haplotypes were related to the increased EC risk (Suppl. Table 2). Furthermore, AT (adjusted OR = 1.42, 95% CI: 1.09–1.85, \( p = 0.010 \)) and GG (adjusted OR = 1.36, 95% CI: 1.05–1.75, \( p = 0.019 \)) haplotypes also conferred higher EC predisposition among subjects with age ≤ 55 years.
MDR Analysis for the Effect of BCL11A SNP-SNP Interaction on EC Risk

MDR analysis was applied to explore the effect of BCL11A SNP-SNP interaction on EC risk. The dendrogram (Figure 2A) and Fruchterman Reingold (Figure 2B) showed that the BCL11A SNP-SNP interaction had a strong redundant effect. The best single-locus and multi-locus models of BCL11A SNPs for EC susceptibility were summarized in Table 3.

**Table 3 The Effect of BCL11A SNPs on the Susceptibility to Endometrial Cancer**

| SNPs ID | Model       | Genotype | Case | Control | Crude Analysis | Adjusted by Age |
|---------|-------------|----------|------|---------|----------------|-----------------|
|         |             |          |      |         | OR (95% CI)    | OR (95% CI)     |                  |
|         |             |          |      |         | p              | p               |                  |
| rs7581162 | Genotype   | AA       | 261  | 288     | 1.16 (0.90–1.50) | 0.259 | 1.16 (0.90–1.50) | 0.257 |
|         | Dominant    | AT       | 204  | 194     | 2.11 (1.24–3.59) | 0.006* | 2.16 (1.27–3.69) | 0.005* |
|         |            | TT       | 204  | 44      | 1.26 (0.98–1.62) | 0.066 | 1.27 (0.99–1.62) | 0.063 |
|         | Recessive   | AA-AT    | 248  | 288     | 1.98 (1.18–3.34) | 0.010 | 2.03 (1.20–3.42) | 0.008* |
|         |            | TT       | 248  | 44      | 1.30 (1.06–1.59) | 0.011 | 1.30 (1.07–1.59) | 0.010 |
| rs10189857 | Genotype   | GG       | 280  | 305     | 1.18 (0.91–1.54) | 0.209 | 1.18 (0.91–1.54) | 0.206 |
|         | Dominant    | GA       | 189  | 174     | 1.74 (1.03–2.95) | 0.038 | 1.78 (1.05–3.01) | 0.033 |
|         |            | AA       | 40   | 25      | 1.25 (0.98–1.61) | 0.076 | 1.26 (0.98–1.62) | 0.072 |
|         | Recessive   | GG-GA    | 469  | 479     | 1.98 (1.18–3.34) | 0.010 | 2.03 (1.20–3.42) | 0.008* |
|         |            | AA       | 40   | 25      | 1.30 (1.06–1.59) | 0.011 | 1.30 (1.07–1.59) | 0.010 |
| rs1427407 | Genotype   | GG       | 291  | 325     | 1.33 (1.02–1.74) | 0.034 | 1.33 (1.02–1.74) | 0.033 |
|         | Dominant    | GT       | 186  | 156     | 1.51 (0.89–2.72) | 0.122 | 1.55 (0.91–2.78) | 0.107 |
|         |            | TT       | 32   | 23      | 1.36 (1.06–1.75) | 0.017 | 1.37 (1.06–1.76) | 0.016 |
|         | Recessive   | GG-GT    | 218  | 179     | 1.40 (0.81–2.43) | 0.228 | 1.43 (0.82–2.49) | 0.204 |
|         |            | TT       | 218  | 32      | 1.29 (1.05–1.59) | 0.016 | 1.30 (1.05–1.60) | 0.014 |
| rs766432  | Genotype   | AA       | 281  | 313     | 1.28 (0.99–1.67) | 0.061 | 1.29 (0.99–1.67) | 0.058 |
|         | Dominant    | AC       | 197  | 171     | 1.57 (0.89–2.77) | 0.121 | 1.60 (0.90–2.84) | 0.107 |
|         |            | CC       | 31   | 22      | 1.32 (1.02–1.69) | 0.032 | 1.32 (1.03–1.70) | 0.029 |
|         | Recessive   | AA-CC    | 228  | 193     | 1.43 (0.81–2.50) | 0.214 | 1.45 (0.83–2.55) | 0.194 |
|         |            | CC       | 31   | 22      | 1.27 (1.03–1.56) | 0.025 | 1.28 (1.04–1.57) | 0.022 |
| rs6729815 | Genotype   | CC       | 283  | 315     | 1.23 (0.95–1.61) | 0.120 | 1.23 (0.95–1.61) | 0.119 |
|         | Dominant    | CT       | 185  | 167     | 1.90 (1.12–3.23) | 0.017 | 1.94 (1.14–3.29) | 0.015 |
|         |            | TT       | 41   | 24      | 1.32 (1.03–1.69) | 0.031 | 1.32 (1.03–1.70) | 0.030 |
|         | Recessive   | CC-CT    | 226  | 191     | 1.76 (1.05–2.96) | 0.033 | 1.79 (1.06–3.02) | 0.028 |
|         |            | TT       | 226  | 41      | 1.30 (1.07–1.60) | 0.010 | 1.31 (1.07–1.61) | 0.009* |

Notes: p values were calculated by logistic regression analysis without or with adjustments for age. Bold indicated that p < 0.05 meant the data was statistically significant. *p indicate that after Bonferroni correction (p < 0.05/5) means the data is statistically significant.

Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.
### Table 4: Stratification Analysis by Age and BMI for the Effect of BCL11A SNPs on the Susceptibility to Endometrial Cancer

| SNPs ID | Model | Genotype | Case | Control | OR (95% CI) | p     | Case | Control | OR (95% CI) | p     |
|---------|-------|----------|------|---------|-------------|-------|------|---------|-------------|-------|
|         |       |          | > 55 years | ≤ 55 years |            |       |      |         |            |       |
| rs7581162 | Allele | A        | 333  | 326     | 1           | 393   | 444  | 1       | 1.40 (1.09–1.82) | 0.010 |
|         |        | T        | 113  | 96      | 1.15 (0.84–1.57) | 0.373 | 179  | 144     | 1.40 (1.09–1.82) | 0.010 |
|         | Genotype | AA       | 117  | 123     | 1           | 144   | 165  | 1       | 1.06 (0.75–1.50) | 0.746 |
|         |        | AT       | 99   | 80      | 1.30 (0.88–1.92) | 0.185 | 105  | 114     | 1.30 (0.88–1.92) | 0.185 |
|         |        | TT       | 7    | 8       | 0.92 (0.32–2.61) | 0.872 | 37   | 15      | 2.84 (1.50–5.40) | 0.001* |
|         | Dominant | AA       | 117  | 123     | 1           | 144   | 165  | 1       | 1.27 (0.91–1.76) | 0.158 |
|         |        | AT-TT    | 106  | 88      | 1.27 (0.87–1.85) | 0.224 | 142  | 129     | 1.27 (0.91–1.76) | 0.158 |
|         |        | TT       | 7    | 8       | 0.82 (0.29–2.30) | 0.707 | 37   | 15      | 2.78 (1.49–5.18) | 0.001* |
|         | Log-additive |        |      |         | 1.18 (0.84–1.64) | 0.343 |      |         | 1.18 (0.84–1.64) | 0.343 |
| rs10189857 | Allele | G        | 344  | 331     | 1           | 405   | 453  | 1       | 1.38 (1.06–1.80) | 0.016 |
|         |        | A        | 102  | 89      | 1.10 (0.80–1.52) | 0.551 | 167  | 135     | 1.38 (1.06–1.80) | 0.016 |
|         | Genotype | GG       | 128  | 131     | 1           | 152   | 174  | 1       | 1.10 (0.78–1.56) | 0.588 |
|         |        | GA       | 88   | 69      | 1.31 (0.88–1.94) | 0.190 | 101  | 105     | 1.31 (0.88–1.94) | 0.190 |
|         |        | AA       | 7    | 10      | 0.72 (0.26–1.94) | 0.512 | 33   | 15      | 2.53 (1.33–4.85) | 0.005* |
|         | Dominant | GG       | 128  | 131     | 1           | 152   | 174  | 1       | 1.28 (0.92–1.78) | 0.142 |
|         |        | GA-AA    | 95   | 79      | 1.23 (0.84–1.81) | 0.291 | 134  | 120     | 1.28 (0.92–1.78) | 0.142 |
|         |        | AA-GA    | 216  | 200     | 1           | 253   | 279  | 1       | 2.44 (1.29–4.60) | 0.006* |
|         | Log-additive |        |      |         | 1.11 (0.80–1.55) | 0.540 |      |         | 1.11 (0.80–1.55) | 0.540 |
| rs1427407 | Allele | G        | 353  | 342     | 1           | 415   | 464  | 1       | 1.44 (1.10–1.89) | 0.008* |
|         |        | T        | 93   | 80      | 1.13 (0.81–1.57) | 0.485 | 157  | 122     | 1.44 (1.10–1.89) | 0.008* |
|         | Genotype | GG       | 134  | 140     | 1           | 157   | 185  | 1       | 1.27 (0.89–1.80) | 0.187 |
|         |        | GT       | 85   | 62      | 1.43 (0.96–2.15) | 0.081 | 101  | 94      | 1.43 (0.96–2.15) | 0.081 |
|         |        | TT       | 4    | 9       | 0.46 (0.14–1.54) | 0.209 | 28   | 14      | 2.37 (1.21–4.67) | 0.012 |
|         | Dominant | GG       | 134  | 140     | 1           | 157   | 185  | 1       | 1.41 (1.01–1.97) | 0.043 |
|         |        | GT-TT    | 89   | 71      | 1.31 (0.89–1.94) | 0.177 | 129  | 108     | 1.31 (0.89–1.94) | 0.177 |
|         |        | TT       | 4    | 9       | 0.41 (0.12–1.35) | 0.142 | 28   | 14      | 2.18 (1.12–4.23) | 0.022 |
|         | Log-additive |        |      |         | 1.14 (0.80–1.60) | 0.472 |      |         | 1.14 (0.80–1.60) | 0.472 |

(Continued)
Table 4 (Continued).

| SNPs ID   | Model   | Genotype | Case | Control | OR (95% CI) | p     | Case | Control | OR (95% CI) | p     |
|-----------|---------|----------|------|---------|-------------|-------|------|---------|-------------|-------|
| rs766432  | Allele  | A        | 350  | 338     | 1           |       | 409  | 459     | 1.42 (1.09–1.85) | 0.010 |
|           |         | C        | 96   | 86      | 1.08 (0.78–1.50) | 0.653 | 163  | 129     | 1.23 (0.87–1.74) | 0.248 |
|           | Genotype| AA       | 130  | 135     | 1           |       | 151  | 178     | 1.23 (1.27–5.09) | 0.008*|
|           |         | AC       | 90   | 68      | 1.37 (0.92–2.04) | 0.118 | 107  | 103     | 1.37 (0.99–1.91) | 0.059 |
|           |         | CC       | 3    | 9       | 0.34 (0.09–1.30) | 0.114 | 28   | 13      | 1.37 (0.99–1.91) | 0.059 |
|           | Dominant| AA-AC    | 130  | 135     | 1           |       | 151  | 178     | 1.23 (1.27–5.09) | 0.008*|
|           |         | CC       | 3    | 9       | 0.30 (0.08–1.14) | 0.078 | 28   | 13      | 2.35 (1.19–4.64) | 0.014 |
|           | Recessive| CC-CT    | 3    | 9       | 0.30 (0.08–1.14) | 0.078 | 28   | 13      | 2.35 (1.19–4.64) | 0.014 |
|           |         | Log-additive | —   | —       | 1.09 (0.77–1.53) | 0.644 | —    | —       | 1.40 (1.08–1.83) | 0.012 |
| rs6729815 | Allele  | C        | 347  | 338     | 1           |       | 404  | 459     | 1.48 (1.14–1.93) | 0.004*|
|           |         | T        | 99   | 86      | 1.12 (0.81–1.55) | 0.490 | 168  | 129     | 1.23 (0.84–1.70) | 0.316 |
|           | Genotype| CC       | 131  | 135     | 1           |       | 152  | 180     | 2.70 (1.42–5.14) | 0.003*|
|           |         | CT       | 85   | 68      | 1.29 (0.87–1.92) | 0.212 | 100  | 99      | 1.39 (1.00–1.94) | 0.050 |
|           |         | TT       | 7    | 9       | 0.80 (0.29–2.21) | 0.668 | 34   | 15      | 2.52 (1.34–4.74) | 0.004*|
|           | Dominant| CC       | 131  | 135     | 1           |       | 152  | 180     | 1.43 (1.11–1.85) | 0.006*|
|           |         | CT-TT    | 92   | 77      | 1.23 (0.84–1.81) | 0.289 | 134  | 114     | 1.43 (1.11–1.85) | 0.006*|
|           | Recessive| CC-CT    | 105  | 105     | 1           |       | 152  | 180     | 1.43 (1.11–1.85) | 0.006*|
|           |         | TT       | 7    | 9       | 0.73 (0.27–2.00) | 0.540 | 34   | 15      | 1.52 (1.34–4.74) | 0.004*|
|           |         | Log-additive | —   | —       | 1.13 (0.81–1.58) | 0.477 | —    | —       | 1.43 (1.11–1.85) | 0.006*|

BMI

| SNPs ID   | Model   | Genotype | Case | Control | OR (95% CI) | p     | Case | Control | OR (95% CI) | p     |
|-----------|---------|----------|------|---------|-------------|-------|------|---------|-------------|-------|
| rs7511162 | Allele  | A        | 293  | 287     | 1           |       | 433  | 483     | 1.42 (1.09–1.83) | 0.006*|
|           |         | T        | 105  | 93      | 1.11 (0.80–1.53) | 0.541 | 187  | 147     | 1.23 (0.85–1.85) | 0.092 |
|           | Genotype| AA       | 109  | 105     | 1           |       | 152  | 183     | 1.33 (0.95–1.85) | 0.092 |
|           |         | AT       | 75   | 77      | 0.94 (0.62–1.42) | 0.756 | 129  | 117     | 2.41 (1.24–4.68) | 0.009*|
|           |         | TT       | 15   | 8       | 1.77 (0.71–4.37) | 0.218 | 29   | 15      | 2.14 (1.12–4.08) | 0.022 |
|           | Dominant| AA-AT    | 109  | 105     | 1           |       | 152  | 183     | 1.45 (1.06–1.99) | 0.022 |
|           |         | TT       | 15   | 8       | 1.82 (0.75–4.41) | 0.188 | 29   | 15      | 1.44 (1.11–1.86) | 0.005*|
| Allele   | G    | 296  | 289  | 1   | 453  | 495  | 1   | 1.37 (1.06–1.78) | 0.017 |
|---------|------|------|------|-----|------|------|-----|-----------------|-------|
| Genotype | GG   | 114  | 108  | 1   | 166  | 197  | 1   | 1.43 (1.02–2.00) | 0.037 |
|         | GA   | 68   | 73   | 0.88 (0.58–1.35) | 0.564 | 121  | 101 | 1.75 (0.89–3.42) | 0.104 |
|         | AA   | 17   | 9    | 1.76 (0.75–4.14) | 0.196 | 23   | 16  |                |       |
| Dominant | GG   | 114  | 108  | 1   | 166  | 197  | 1   |                |       |
|         | GA-AA | 85   | 82   | 0.98 (0.65–1.46) | 0.917 | 144  | 117 |                |       |
| Recessive | GG-GA | 182  | 181  | 1   | 287  | 298  | 1   |                |       |
|         | AA   | 17   | 9    | 1.85 (0.80–4.28) | 0.153 | 23   | 16  |                |       |
| Log-additive | —    | —    | —    | —   | —    | —    | —   | 1.38 (1.06–1.78) | 0.016 |

| Allele   | T    | 94   | 81   | 1.14 (0.81–1.60) | 0.442 | 156  | 121 | 1.41 (1.08–1.84) | 0.012 |
| Genotype | GG   | 118  | 117  | 1   | 173  | 208  | 1   | 1.57 (1.12–2.21) | 0.009*|
|         | GT   | 68   | 65   | 1.04 (0.68–1.59) | 0.867 | 118  | 91  | 1.57 (0.77–3.18) | 0.214 |
|         | TT   | 13   | 8    | 1.56 (0.62–3.95) | 0.344 | 19   | 15  |                |       |
| Dominant | GG   | 118  | 117  | 1   | 173  | 208  | 1   |                |       |
|         | GT-TT | 81   | 73   | 1.09 (0.73–1.64) | 0.664 | 137  | 106 | 1.57 (1.13–2.17) | 0.007*|
| Recessive | GG-GT | 186  | 182  | 1   | 291  | 299  | 1   |                |       |
|         | TT   | 13   | 8    | 1.54 (0.62–3.85) | 0.352 | 19   | 15  | 1.33 (0.66–2.68) | 0.419 |
| Log-additive | —    | —    | —    | —   | —    | —    | —   | 1.41 (1.08–1.84) | 0.012 |

| Allele   | A    | 303  | 298  | 1   | 464  | 507  | 1   | 1.37 (1.05–1.78) | 0.019 |
| Genotype | AA   | 71   | 70   | 1   | 165  | 199  | 1   | 1.52 (1.09–2.12) | 0.014 |
|         | AC   | 12   | 7    | 1.00 (0.65–1.52) | 0.985 | 126  | 101 | 1.58 (0.78–3.21) | 0.208 |
|         | CC   | 116  | 114  | 1.64 (0.62–4.34) | 0.321 | 19   | 15  |                |       |
| Dominant | AA   | 83   | 77   | 1   | 165  | 199  | 1   |                |       |
|         | AC-CC | 187  | 184  | 1.05 (0.70–1.58) | 0.799 | 145  | 116 | 1.53 (1.11–2.10) | 0.010 |
| Recessive | AA-AC | 12   | 7    | 1   | 291  | 300  | 1   |                |       |
|         | CC   | —    | —    | —   | —    | —    | —   | 1.39 (1.07–1.82) | 0.015 |
| Log-additive | —    | —    | —    | —   | —    | —    | —   |                |       |
Table 4 (Continued).

| SNPs ID  | Model    | Genotype | Case | Control | OR (95% CI)   | p   | Case | Control | OR (95% CI)   | p   |
|----------|----------|----------|------|---------|---------------|-----|------|---------|---------------|-----|
| rs6729815 | Allele   | C        | 300  | 193     | 1             | 451 | 504  | 1.50 (1.15–1.95) | 0.003*  |
|          |          | T        | 98   | 89      | 1.08 (0.77–1.49) | 0.665 | 169 | 126     | 1.50 (1.15–1.95) | 0.003*  |
|          | Genotype | CC       | 117  | 110     | 1             | 166 | 205  | 1.49 (1.15–1.93) | 0.003*  |
|          |          | CT       | 66   | 73      | 0.85 (0.56–1.30) | 0.456 | 119 | 94      | 1.57 (1.12–2.21) | 0.009*  |
|          |          | TT       | 16   | 8       | 1.84 (0.76–4.50) | 0.179 | 25  | 16      | 1.99 (1.03–3.87) | 0.042  |
|          | Dominant | CC       | 117  | 110     | 1             | 166 | 205  | 1.49 (1.15–1.93) | 0.003*  |
|          |          | CT-TT    | 82   | 81      | 0.95 (0.63–1.42) | 0.802 | 144 | 110     | 1.63 (1.18–2.26) | 0.003*  |
|          | Recessive| CC-CT    | 183  | 183     | 1             | 285 | 299  | 1.68 (0.88–3.23) | 0.116  |
|          |          | TT       | 16   | 8       | 1.96 (0.81–4.72) | 0.134 | 25  | 16      | 1.49 (1.15–1.93) | 0.003*  |
|          | Log-additive | —     | 107  | —       | —             | —   | —    | —       | —             | —   |

Notes: p values were calculated by Fisher’s exact test or logistic regression analysis with adjustments for age. Bold indicated that p < 0.05 meant the data was statistically significant. *p indicate that after Bonferroni correction (p < 0.05/5) means the data is statistically significant.

Abbreviations: SNP, single nucleotide polymorphism; BMI, body mass index; OR, odds ratio; 95% CI, 95% confidence interval.
in Table 5. In the single-locus model, rs1427407 was the most influential attributor for EC risk (testing accuracy = 0.5356, cross-validation consistency (CVC) = 10/10). In the multi-locus model, the best combination was a two-locus model containing rs7581162 and rs766432 (testing accuracy= 0.5336, CVC= 5/10).

FPRP Analysis for the Significant Findings

FPRP analysis was carried out to interrogate whether the significant findings were deserving attention (Suppl_Table 4). At the prior probability level of 0.1, the significant association for rs7581162, rs10189857, rs1427407, rs766432 and rs6729815 remained noteworthy in the overall analysis (FPRP < 0.2) and statistical power > 85% under the allele model. In the subgroup at age < 55 years, significant findings remained noteworthy for these variants.

The Relationship of BCL11A SNPs with Characteristics Among EC Patients/Healthy Controls

The relationship of BCL11A SNPs with CEA, AFP, CA199, and CA125 among EC patients/healthy controls were assessed, as displayed in Suppl_Table 5. However, no statistically association was observed.

Discussion

In this study, we found that rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 in BCL11A were associated with EC susceptibility in Chinese Han women, especially in subjects aged ≤ 55 years and BMI <24 kg/m². Haplotype analysis showed that rs10189857 and rs1427407 were in linkage disequilibrium, and the A<sub>rs10189857</sub>T<sub>rs1427407</sub> and G<sub>rs10189857</sub>G<sub>rs1427407</sub> haplotypes were related to the increased EC risk. MDR analysis indicated that rs1427407 was the most influential attributor on EC risk in the single-locus model, and the best combination was a two-locus model containing rs7581162 and rs766432. Our study is the first to report that BCL11A variants were risk factors for EC in Chinese Han females.
The BCL11A gene, located on 2p16, spans 102 kb, which is considered to be an oncogene in malignant haematological diseases, and is first detected in B-cell chronic lymphocytic leukaemia.\(^{21}\) BCL11A can cause transcriptional repression of mammalian target genes by binging to DNA motifs and promoting the deacetylation of H3/H4 histone.\(^{22}\) BCL11A may participate in cell cycle and cell growth by inhibiting of P21 induction.\(^{23}\) In addition, BCL11A is involved in cell apoptosis by upregulating the expression of BCL2, BCL2-xL, and MDM2 and inhibiting the activity of P53.\(^{24}\) These studies support the possible involvment of BCL11A in tumorigenesis. Studies have found that BCL11A is abnormally expressed in EC tissues, which is related to EC classification, lymph node metastasis, tumor differentiation, histological type, ER/PR expression.\(^{13}\) Targeted next-generation sequencing evaluated BCL11A mutations in patients with endometriosis.\(^{14}\) These lines of evidence have led us to formulate the hypothesis that BCL11A could be of pathogenic importance in EC.

Genetic variations of BCL11A gene may be associated with gene expression, thereby affecting the occurrence and development of disease. To date, BCL11A polymorphisms have been reported to be associated with various cancer, such as pancreatic cancers, laryngeal squamous cell carcinoma, and chronic lymphocytic leukemia.\(^{25–27}\) Previous study have revealed that BCL11A rs7579014 is a risk locus for EC occurrence and is related to the expression of BCL11A in endometrial tumors.\(^{15}\) BCL11A rs148261157 has been reported to increase the risk of EC.\(^{16}\) However, the association between rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 in BCL11A and EC susceptibility has not been

---

![Figure 2](https://doi.org/10.2147/PGPM.S345772)

**Figure 2.** The dendrogram (A) and Fruchterman Reingold (B) of BCL11A SNP-SNP interaction for EC risk. (A) Short connections among nodes represent stronger redundant interactions. Green and blue color indicated weak interactions. (B) This graphical model describes the percent entropy explained by each SNP. Values in nodes represent the information gains of individual attribute (main effects). Values between nodes are information gains of each pair of attributes (interaction effects). Positive percent entropy indicates synergy whereas the negative percent entropy indicates redundancy.

---

**Table 5.** SNP–SNP Interaction Models of the BCL11A Gene the Predisposition of Endometrial Cancer

| Model                              | Training Bal. Acc. | Testing Bal. Acc. | CVC | OR (95% CI)       | p   |
|------------------------------------|--------------------|------------------|-----|------------------|-----|
| rs1427407                          | 0.5356             | 0.5356           | 10/10 | 1.35 (1.05–1.74) | 0.0277 |
| rs7581162, rs766432                | 0.5412             | 0.5336           | 5/10  | 1.42 (1.08–1.86) | 0.0106 |
| rs7581162, rs10189857, rs766432     | 0.5447             | 0.5287           | 4/10  | 1.44 (1.12–1.86) | 0.0047 |
| rs7581162, rs10189857, rs1427407, rs766432 | 0.5470         | 0.5208           | 4/10  | 1.47 (1.14–1.90) | 0.0029 |
| rs7581162, rs10189857, rs1427407, rs766432, rs6729815 | 0.5479        | 0.5178           | 10/10 | 1.46 (1.14–1.88) | 0.0029 |

Notes: p values were calculated using \(\chi^2\) tests. Bold indicated that \(p < 0.05\) meant the data was statistically significant.

Abbreviations: MDR, multifactor dimensionality reduction; Bal. Acc., balanced accuracy; CVC, cross-validation consistency; OR, odds ratio; CI, confidence interval.
reported. Our study displayed that the MAFs of these five SNPs in BCL11A were higher in EC patients than in healthy controls, and had the increased risk effect on EC predisposition. To date, the functions of these SNPs have not been reported. In bioinformatics analysis, results from HaploReg v4.1 database displayed that these SNPs may be associated with promoter/Enhancer histone marks, DNAse, proteins bound, motifs changed, NHGRI/EBI GWAS hits, and/or GRASP QTL hits.28 Previously, rs766432 in the intronic regions of BCL11A gene may affect the binding of protein to this region.29 Rs1427407 in the second intron of the trans-acting element BCL11A is associated with the expression of BCL11A.30 Based on GTEx Portal database, the genotypes of rs1427407 was related to the mRNA expression of BCL11A in cell. These results suggest that these loci may be involved in EC carcinogenic by affecting the expression or function of BCL11A, which requires further experimental confirmation.

Age is a risk factor for the development of EC.31 A study reported that median age at diagnosis of endometrial cancer was 55 years.32 Besides, the mean ages of EC patients and the controls were respectively 54.94 ± 8.85 years and 54.61 ± 9.07 years in the study. In order to explore the contribution of age, we have divided the cases and controls into two groups as ≤ 55 years and > 55 years. The results of age-stratified analysis displayed that rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 were associated with EC susceptibility in subjects aged ≤ 55 years, but not in subjects aged > 55 years, suggesting that the risk association between BCL11A variants and EC predisposition might be age-dependent. Obesity is a risk factor for endometrial cancer risk and mortality.33 Using body mass index (BMI) as a measure of obesity, we assessed the association of BCL11A variants with BMI. Among those aged 55 or younger and those had BMI < 24 kg/m², we found that rs7581162-T, rs10189857-A, rs1427407-T, rs766432-C, and rs6729815-T were risk factors for the development of EC. The underlying mechanism of this correlation awaits further study. Haplotype-based analysis may be more effective than single-locus analysis when there is linkage disequilibrium among SNPs.34 Haplotype analysis revealed that rs10189857 and rs1427407 were in linkage disequilibrium, and the A<sub>rs10189857</sub>T<sub>rs1427407</sub> and G<sub>rs10189857</sub>G<sub>rs1427407</sub> haplotypes were related to an increased risk of EC. MDR analysis was used to identify specific combination effects of genetic variants with high-order interactions.35 In this study, we found that rs1427407 was the most influential attributor for EC risk in the single-locus model, and the best combination was the two-locus model incorporating rs7581162 and rs766432 in BCL11A.

Although our results provided evidence on the relationship between BCL11A polymorphisms and EC predisposition, several limitations should not be neglected. First, our subjects were recruited from a hospital that may contribute to selection bias. Second, due to the lack of adequate personal information, this study failed to assess the impact of gene-environment interactions on EC susceptibility and the association of genetic variants with clinicopathological data of EC patients. In the future, we would like to enlarge sample size and complete the clinicopathological data to evaluate the relationship. Third, although our findings suggested that BCL11A variants were associated with increased the risk of EC, the potential mechanisms and functions of these SNPs underlying the association have not been revealed. Fourth, the expression data of BCL11A is missing. In subsequent research, we would design detailed experiments to further explore the expression data of BCL11A and the potential mechanisms and functions of these SNPs in EC.

In conclusion, our study first provides evidence that rs7581162, rs10189857, rs1427407, rs766432, and rs6729815 in BCL11A are risk factors for EC occurrence in Chinese Han women. These findings increased our understanding of the role of BCL11A gene in EC pathogenesis. However, the expression data of BCL11A and the association between SNPs and BCL11A expression need further explore in the more detailed experiments.

Acknowledgments
We are grateful to the subjects for their participation in the study and to the hospital staff for their contributions to the study.

Funding
This study was funded by Hainan Province Science and Technology Special Fund (ZDYF2019123 and ZDYF2020131) and Hainan Province Clinical Medical Center.
Disclosure

All authors declare that they have no competing interests in this work.

References

1. Cavusoglu M, Sozmen Citiz D, Orsoy A, et al. Diffusion-weighted MRI of postmenopausal women with vaginal bleeding and endometrial thickening: differentiation of benign and malignant lesions. J Belg Soc Radiol. 2016;109(1):70. doi:10.5334/jbr.1118
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. doi:10.3322/caac.21492
3. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115–132. doi:10.3322/caac.21338
4. Li X, Zheng S, Chen S, Qin F, Lau S, Chen Q. Trends in gynaecological cancers in the largest obstetrics and gynaecology hospital in China from 2003 to 2013. Tumour Biol. 2015;36(7):4961–4966. doi:10.1007/s13277-015-3143-6
5. Braun MM, Overbeek-Weger EA, Grumbo RJ. Diagnosis and management of endometrial cancer. Am Fam Physician. 2016;93(6):468–474.
6. Du Q, Guo X, Zhang X, et al. SYN2 variant rs9365723 is associated with colorectal cancer risk in Chinese Han population. Int J Biol Markers. 2016;31(2):138–143. doi:10.5301/ijb.2016.00182
7. He N, Liu L, Duan X, et al. Identification of a shared protective genetic susceptibility locus for colorectal cancer and gastric cancer. Tumour Biol. 2016;37(2):2443–2448. doi:10.1007/s13277-015-4070-2
8. Jin TB, Du S, Zhu XK, et al. Polymorphism in the IL4R gene and clinical features are associated with glioma prognosis: analyses of case-cohort studies. Medicine. 2016;95(31):e4231. doi:10.1097/MD.0000000000004231
9. Lazarus KA, Hadi F, Zambon E, et al. BCL11A interacts with SOX2 to control the expression of epigenetic regulators in lung squamous carcinoma. Nat Commun. 2018;9(1):3327. doi:10.1038/s41467-018-05790-5
10. Huang HT, Chen SM, Pan LB, Yao J, Ma HT. Loss of function of SWI/SNF chromatin remodeling genes leads to genome instability of human lung cancer. Oncol Rep. 2015;33(1):283–291. doi:10.3892/or.2014.3584
11. Zhu L, Pan R, Zhou D, Ye G, Tan W. BCL11A enhances stemness and promotes progression by activating Wnt/β-catenin signaling in breast cancer. Cancer Manag Res. 2019;11:2997–3007. doi:10.2147/CMAR.S199368
12. Park SY, Lee SJ, Cho HJ, et al. Epsilon-globin HBE1 enhances radiotherapy resistance by down-regulating BCL11A in colorectal cancer cells. Cancers. 2019;11(4):498. doi:10.3390/cancers11040498
13. You D, Lin L, Wang Q, et al. Clinical significance of BCL11A expression in ER-negative and PR-negative endometrial carcinoma. Int J Clin Exp Pathol. 2018;11(6):3068–3075.
14. Er TK, Su YF, Wu CC, et al. Targeted next-generation sequencing for molecular diagnosis of endometriosis-associated ovarian cancer. J Mol Med. 2016;94(7):835–847. doi:10.1007/s00109-016-1395-2
15. O’Mara TA, Spurkle AB, Glubb DM. Analysis of promoter-associated chromatin interactions reveals biologically relevant candidate target genes at endometrial cancer risk loci. Cancers. 2019;11(10). doi:10.3390/cancers11101440
16. O’Mara TA, Glubb DM, Amant F, et al. Identification of nine new susceptibility loci for endometrial cancer. Nat Commun. 2018;9(1):3166. doi:10.1038/s41467-018-05427-7
17. Oeth P, Del Mistro G, Marnellos G, Shi T, van den Boom D. Qualitative and quantitative genotyping using single base primer extension coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MassARRAY). Methods Mol Biol. 2009;578:307–343.
18. Ellis JA, Ong B. The MassARRAY(R) system for targeted SNP genotyping. Methods Mol Biol. 2017;1492:77–94.
19. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst. 2004;96(6):434–442. doi:10.1093/jnci/djh075
20. Zhou Z, Lu H, Zhu J, et al. METTL14 gene polymorphisms confer neuroblastoma susceptibility: an eight-center case-control study. Mol Ther Nucleic Acids. 2020;22:17–26. doi:10.1016/j.omtn.2020.08.009
21. Menega MA, Pulford K, Keskin S, et al. Gains of the proto-oncogene BCL11A and nuclear accumulation of BCL11A(XL) protein are frequent in primary mediastinal B-cell lymphoma. Leukemia. 2006;20(10):1880–1882. doi:10.1038/sj.leu.2404342
22. Chen Z, Luo HY, Steinberg MH, Chui DH. BCL11A represses HBG transcription in K562 cells. Blood Cells Mol Dis. 2009;42(2):144–149. doi:10.1016/j.bcmd.2008.12.003
23. Yin B, Delwel R, Valk PJ, et al. A retroviral mutagenesis screen reveals strong cooperation between Bcl11a overexpression and loss of the Nf1 tumor suppressor gene. Blood. 2009;113(5):1075–1085. doi:10.1182/blood-2008-03-144436
24. Yu Y, Wang J, Khalid W, et al. Bcl11a is essential for lymphoid development and negatively regulates p53. J Exp Med. 2012;209(13):2467–2483. doi:10.1084/jem.20121846
25. Pierce BL, Austin MA, Ahsan H. Association study of type 2 diabetes genetic susceptibility variants and risk of pancreatic cancer: an analysis of PanScan-1 data. Cancer Causes Control. 2011;22(6):877–883. doi:10.1007/s10552-011-9760-5
26. Zhou J, Yang Y, Zhang D, Zhou L, Tao L, Lu LM. Genetic polymorphisms and plasma levels of BCL11A contribute to the development of laryngeal squamous cell carcinoma. PLoS One. 2017;12(2):e0171116. doi:10.1371/journal.pone.0171116
27. Pfeifer D, Pantic M, Skatulla I, et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood. 2007;109(3):1202–1207. doi:10.1182/blood-2006-07-032452
28. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40(Database issue):D930–934. doi:10.1093/nar/gkr917
29. Neishabury M, Zamani F, Keyhani E, et al. The influence of the BCL11A polymorphism on the phenotype of patients with beta thalassemia could be affected by the beta globin locus control region and/or the Xmn1-HBG2 genotypic background. Blood Cells Mol Dis. 2013;51(2):80–84. doi:10.1016/j.bcmd.2013.02.007
30. Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science. 2013;342(6155):253–257. doi:10.1126/science.1242088
31. Zhou JY, Zhang L, Wei LH, Wang JL. Endometrial carcinoma-related genetic factors: application to research and clinical practice in China. BJOG. 2016;123(Suppl 3):90–96. doi:10.1111/1471-0528.14007
32. Nitecki R, Fu S, Lefkowits C, et al. Employment disruption following the diagnosis of endometrial cancer. *Gynecol Oncol*. 2021;160(1):199–205. doi:10.1016/j.ygyno.2020.10.041

33. Shaw E, Farris M, McNeil J, Friedenreich C. Obesity and endometrial cancer. *Recent Results Cancer Res*. 2016;208:107–136. doi:10.1007/978-3-319-42542-9_7

34. Niu T, Qin ZS, Xu X, Liu JS. Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *Am J Hum Genet*. 2002;70(1):157–169. doi:10.1086/338446

35. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics*. 2003;19(3):376–382. doi:10.1093/bioinformatics/btf869