Cancer
Genetics

Long noncoding RNA *CCAT1* polymorphisms are associated with the risk of colorectal cancer

Yingjun Li[a], Fangyuan Jing[a], Ye Ding[a], Qingfang He[b], Yaohong Zhong[a], Chunhong Fan[a,*]

[a] *Department of Public Health, Hangzhou Medical College, Hangzhou, Zhejiang, China;* [b] *Department of Chronic Non-Communicable Diseases Control and Prevention, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, Zhejiang, China*

Abstract

Colorectal cancer associated transcript 1 (*CCAT1*) is a novel long noncoding RNA, whose overexpression is evident in both early phase of tumorigenesis and later disease stages in colorectal cancer (CRC). No study has explored the relationship between *CCAT1* polymorphisms and CRC risk. In the present study, a case-control study was conducted to investigate the association between *CCAT1* polymorphisms and CRC risk in Chinese population. We identified that *CCAT1* rs67085638 polymorphism was associated with an increased risk of CRC (OR = 1.72, 95%CI = 1.14–2.58, $P = 0.009$ in heterozygote codominant model; OR = 1.67, 95%CI = 1.13–2.47, $P = 0.010$ in dominant model). Moreover, *CCAT1* rs7013433 polymorphism was associated with late clinical stage (OR = 1.82, 95%CI = 1.16–2.86, $P = 0.009$ in heterozygote codominant model; OR = 1.72, 95%CI = 1.13–2.63, $P = 0.012$ in dominant model). Our finding proposed a link between *CCAT1* polymorphisms with CRC risk as well as different clinical stages.

Keywords Long noncoding RNA, CCAT1, Polymorphism, Colorectal cancer.

Introduction

Colorectal cancer (CRC) is one the most common types of gastrointestinal cancer and contributes to the third cancer cause and fourth cancer death worldwide in 2012 [1]. There are estimated to be 135,430 individuals newly CRC cases and 50,260 newly CRC deaths in the United States in 2017 [2]. Although the etiology of CRC is complex and remains unknown, many studies have reported the link between single nucleotide polymorphisms (SNPs) in protein-coding genes and functional noncoding RNAs with the risk of CRC development [3–9].

Long noncoding RNAs (lncRNAs) are a class of noncoding RNAs greater than 200 nucleotides in length, which has been reported to play important roles as epigenetic regulators in gene expression [10,11]. Increasing evidence indicates that abnormal expression of lncRNAs is involved in the diversified mechanisms in pathogenesis of various diseases including cancer [12–17]. Polymorphisms of lncRNAs might exert various effects on their expressions and/or functions, thus regulating individual cancer susceptibility [18–20].

Colorectal cancer associated transcript 1 (*CCAT1*) is a novel 11.2 kb lncRNA, whose overexpression is evident in both early phase of tumorigenesis and later disease stages in CRC [21]. Xiang et al. has demonstrated that *CCAT1* plays a role in *MYC* transcriptional regulation and promotes long-range chromatin looping. Knockdown of *CCAT1* reduces long-range interactions between the *MYC* promoter and its enhancers [22]. *MYC* transcription and cell growth are tightly correlated with the presence of *CCAT1* RNA in a variety of tumor types [23]. Moreover, Zhao et al. has found that oncogenic SNP rs6983267 in the *MYC* enhancer region plays a role in the regulation of *CCAT1* expression and are associated with endometrial carcinoma [24]. A meta-analysis has demonstrated that increased expression level of *CCAT1* is associated with clinicopathological features in relevant cancers including TNM stage [25]. Additionally, emerging evidence has shown that *CCAT1* is associated with CRC development and prognosis. Zhao et al. has found that increased plasma
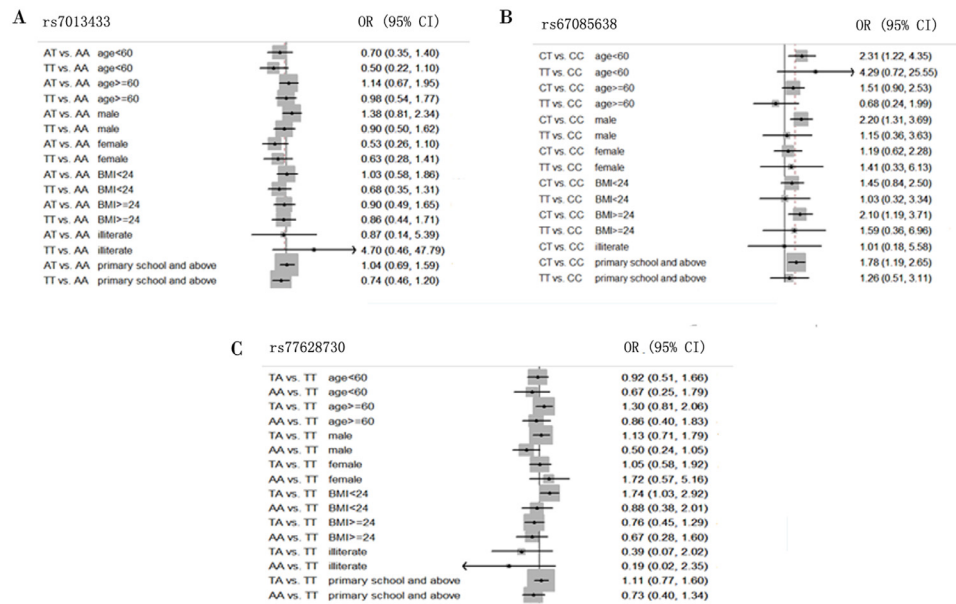
Fig. 1 Forest plots shows odds ratio for the association between *CCAT1* rs7013433 (A), rs67085638 (B), and rs77628730 (C) polymorphisms and the risk of colorectal cancer stratified by age, gender, BMI, and education level in codominant model.

CCAT1 could be used as a predictive biomarker for CRC screening [26]. Ozawa et al. has found that high expression of *CCAT1* is significantly associated with poor recurrence free survival and overall survival in CRC patients [27].

However, to the best of our knowledge, no study has explored the relationship between *CCAT1* polymorphisms and CRC risk. Therefore, a case-control study was conducted to investigate the association between *CCAT1* polymorphisms and CRC risk in Chinese population.

Materials and methods

Study subjects

The CRC cases were recruited from our ongoing CRC study, in which participants were from the First Affiliated Hospital of Zhejiang University and Zhejiang Provincial People's Hospital. Eligible cases were incident and histologically confirmed primary colorectal cancer, mentally competent to complete the interview and with no previous diagnosis of familial adenomatous polyposis, ulcerative colitis or Crohn's disease. The healthy controls were obtained from a community in Hangzhou, Zhejiang Province by random sampling. The research adopted the face-to-face questionnaire survey to obtain the demographic characteristics and lifestyle-related factors of the selected objects. A structured questionnaire was completed after the objects were asked to agree and signed the informed consent. After interview, 5 ml blood of the subject was collected for DNA isolation.

SNP selection and genotyping

Genomic DNA was isolated from peripheral blood samples for each study subject using the modified salting-out proce-

dure [28]. We selected tagSNPs with minor allele frequencies (MAF) > 0.10 in Han Chinese Beijing according to the 1000 Genome Projects. We amplified the region to 2000 bp upstream as promoter region. We also selected tagSNPs from 3′ UTR region as genetic variants in these regions always played functional roles in disease development [29–31]. Thus, three tagSNPs (rs7013433 A $>$ T in promoter region, rs67085638 C $>$ T in 3′ UTR region, and rs77628730 T $>$ A in 3′ UTR region) were selected when linkage disequilibrium (LD) between pair-wise SNPs was with a minimum r^2 of 0.80. Genotyping of included SNPs were conducted by using a custom-by-design 48-Plex SNPscan^TM Kit (Gensky Biotechnology Co., Shanghai, China) in July 2017. This kit was developed based on double ligation and multiplex fluorescence PCR [32]. We randomly selected 5% DNA samples for repeated detection and the concordance rates were $>$99%.

Statistical analysis

The analyses were mainly carried out by SAS 9.2 software (SAS Institute, Cary, NC, USA). Chi-square test was used to test the differences in the distribution of selected demographic variables and genotypes of tagSNPs. A goodness-of-fit chi-square test was used to test the Hardy–Weinberg equilibrium for each SNP. Unconditional logistic regression analyses with odds ratios (ORs) and 95% confidence intervals (CIs) were conducted to estimate the associations between each SNP and CRC susceptibility. Estimates were calculated in crude (Model 1), and additionally adjusted for age ($\leq$44, 45–59, 60–74, $\geq$75), gender (male, female), body mass index (BMI) ($<$18.5, 18.5–23.9, 24–27.9, $\geq$28), education level (illiterate, primary school, middle school and above), red meat ($\geq$4 times/week, 2–3 times/week, $\leq$1 time/week), root vegetable ($\geq$6 times/week, 4–5 times/week, $\leq$3 times/week) (Model 2). Codominant and dominant genetic models were conducted to