Polymorphisms of a novel long non-coding RNA RP11-108K3.2 with colorectal cancer susceptibility and their effects on its expression

Danie Jiang1, Mingjuan Jin1, Ding Ye1,2, Yingjun Li3, Fangyuan Jing3, Xiaocong Zhang1, Qilong Li4 and Kun Chen1,5

Abstract
Background: RP11-108K3.2 was recently identified as a novel long non-coding RNA (lncRNA) transcript, and several single nucleotide polymorphisms (SNPs) have been identified in its coding region. This study aimed to explore the associations of tagSNPs in RP11-108K3.2 with the risk of colorectal cancer and their effects on its expression.

Methods: A total of 821 colorectal cancer cases and 857 healthy controls were enrolled into this two-stage case-control study. Demographic characteristics and lifestyle information were collected by a validated questionnaire. Six tagSNPs were genotyped by using Sequenom MassARRAY platform. A total of 71 additional colorectal cancer cases were recruited, of which the genotypes of potential polymorphisms and the RP11-108K3.2 expression levels were determined.

Results: In the discovery set, only the rs2470151 C/T polymorphism was found to have a promising association with the risk of colorectal cancer, and this polymorphism was further replicated in the validation set with a significantly decreased risk of colorectal cancer (adjusted odds ratio 0.73; 95% confidence interval 0.55, 0.97). Combined discovery set and validation set together, this negative association was found both in the heterozygote codominant model and the dominant model. Furthermore, colorectal cancer patients carrying rs2470151 CT/TT genotypes had a marginally lower RNA expression of RP11-108K3.2 than those carrying the CC genotype. Stratified analyses showed the association between rs2470151 and the risk of colorectal cancer were influenced by family history of cancer, smoking, alcohol consumption, and tea drinking.

Conclusions: These findings suggest that RP11-108K3.2 rs2470151 had a significant association with the risk of colorectal cancer; this may help to predict the susceptibility of colorectal cancer in Chinese populations.

Keywords
Long non-coding RNA, RP11-108K3.2, single nucleotide polymorphism, rs2470151, colorectal cancer

Date received: 25 March 2019; revised: 18 October 2019; accepted: 21 October 2019
**Introduction**

Colorectal cancer (CRC) is the third most common cancer and ranks fourth in cancer deaths globally, accounting for roughly 1.2 million new cases and 600,000 deaths every year. Since CRC is a significant threat to public health, a substantial number of studies have investigated its etiology. Apart from age and sex, the following factors have been identified and established to be associated with the risk of CRC: smoking, excessive alcohol consumption, high consumption of red and processed meat, obesity, and diabetes.

A substantial number of genetic components play important roles in CRC susceptibility. Elevated genome-wide association studies have indicated that single nucleotide polymorphisms (SNPs) were involved. Among the numerous SNPs demonstrated to be related to CRC, variants located on the coding region of long non-coding RNAs (lncRNAs) have drawn attention. For example, a two-stage case-control study suggests that genetic variants of *HOTAIR* rs7958904 may act as a potential biomarker for predicting the risk of CRC.

lncRNAs are a large category of transcripts without protein-coding potential, which are >200 nt in length. Although the definite biological roles of most lncRNAs are still unknown, a growing number of literature suggests that they have a wide variety of roles in chromatin remodeling, transcriptional activation, transcriptional repressor, and RNA degradation. The genetic polymorphisms on lncRNA coding regions may have implications for many human diseases, including cancer, by changing their expression levels.

*RP11-108K3.2* was recently identified as a novel lncRNA, and its coding region maps on the 15q21 chromosome, which is overlapped by the cytochrome P450 family 19 subfamily A member 1 (*CYP19A1*) gene. *CYP19A1* influences estrogen-metabolizing enzymes and may change the CRC risk through an inflammation-related mechanism. A previous population-based, case-control study suggests that several polymorphisms in the *CYP19A1* gene are associated with the risk of CRC, which may have a stronger effect in women due to the influences of sex hormone.

In this study, we hypothesized that tagSNPs located on *RP11-108K3.2* may be involved in the susceptibility to CRC. To test this hypothesis, we conducted a two-stage, case-control study to explore the association between the tagSNPs of *RP11-108K3.2* and the risk of CRC in a Chinese population. We further analyzed the expression levels of *RP11-108K3.2* among different genotype carriers if the polymorphism has been proven to be associated with the risk of CRC.

**Materials and methods**

**Study subjects**

This study was performed in Jiashan County, Zhejiang Province, China, based on a population screening program. All CRC cases were histopathologically confirmed by the Jiashan Institute of Cancer Prevention and Treatment since 1989. Healthy controls were recruited in parallel from the same population and matched to the cases by age (± 5 years), gender, and residential area. Subjects were excluded if they self-reported a history of familial adenomatous polyposis, ulcerative colitis, or Crohn disease. The recruitment details of the study population have been described previously. Briefly, this two-stage, case-control study ultimately enrolled 320 CRC cases and 319 healthy controls in the discovery set; and 501 CRC cases and 538 healthy controls in the validation set. A face-to-face interview was conducted by trained interviewers using a structured questionnaire to collect the information of demographic characteristics (e.g. age, gender, and body mass index (BMI)), family history of cancer and lifestyle factors (e.g. smoking, alcohol consumption, and tea drinking). The definition of a family history of cancer was a first- or second-degree relative having had cancer at any site. Smokers were defined as those who smoked at least one cigarette per day for more than 1 year, or >300 cigarettes in less than 3 months. Alcohol drinkers were defined as those who consumed at least one alcohol drink per day for more than 3 months. Tea drinkers were defined as those who consumed more than two times of tea drink per week for more than 2 months. Blood sample of each participant was collected and stored at −80°C for DNA isolation.

This study also recruited 71 CRC tissues from patients who underwent surgical operations at the Department of Gastrointestinal Surgery, Hangzhou First People’s Hospital from July 2013 to December 2013. The tissues were preserved in the RNA Later® Stabilization Solution (Invitrogen, Carlsbad, CA, USA) and stored at −80°C.

The written informed consent was obtained from all subjects and this study was approved by the Medical Ethical Committee of Zhejiang University School of Medicine.

**SNP selection and genotyping**

We selected the tagSNPs that were located on the 2000bp upstream region of *RP11-108K3.2* with pairwise correlation r^2>0.80 and minor allele frequencies (MAF) > 0.10. Genomic DNA was extracted from peripheral blood by using the modified salting-out procedure. Genotyping was performed by the MassARRAY molecular weight array analysis system (BioMiao Biological Technology Co., Beijing, China). A negative control was included in each 96-well plate for quality control. The 5% of samples were randomly selected for repeated detection and the concordance rate was >99%. Primer information of SNPs of *RP11-108K3.2* is shown in Supplementary Table 1.

**RNA isolation and quantitative RT-PCR**

Total RNA was isolated using a homogenizer (IKA® Works Guangzhou, China) and TRIzol reagent (Invitrogen,
Carlsbad, CA, USA) and then purified using the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. A reverse transcription real-time PCR (RT-PCR) was conducted with StepOnePlus instrument (Applied Biosystems, Foster City, CA, USA) to synthesize cDNA. We performed quantitative real-time PCR in StepOnePlus System (Applied Biosystems, Foster City, CA, USA) using SYBR Green Master Mix (Vazyme, Nanjing, China) to quantify relative RP11-108K3.2 expression levels. The specific forward primer and reverse primer were 5′-GCTGGAACCAATGGATGA-3′ and 5′-CCTGGAGAATGTAGCAATGGAA-3′. The β-actin was selected as the endogenous control. The specific forward primer and reverse primer were 5′-GTGGCG CGAGGACTTTGATTG-3′ and 5′-CCTGGAGAATGTAGCAATGGAA-3′. All procedures were carried out in triplicate. For quantitative results, the expression of RP11-108K3.2 was calculated according to fold change using the $2^{-\Delta\Delta C_{t}}$ method.

**Statistical analysis**

The Student’s t-test and Pearson’s $\chi^{2}$ test were applied to evaluate the differences of characteristics between cases and controls. Hardy–Weinberg equilibrium was assessed by goodness-of-fit $\chi^{2}$ test. The association between tagSNPs and the risk of CRC was estimated by calculating the adjusted odds ratios (ORs) and corresponding 95% confidence intervals (CIs) in multivariate logistic regression model. Those tagSNPs with a $P$-value less than 0.20 in the association study of discovery set were further analyzed in the validation set. Pooled analyses were also conducted to estimate the combined effect of these two sets. Different expression levels of lncRNA in samples of different genotypes were analyzed using ANOVA analysis or t-test.

All analyses were performed with SAS software (version 9.2; SAS Institute, Inc., Cary, NC, USA) with two-sided $P$-values. $P < 0.05$ was considered statistically significant.

**Results**

**Characteristics of the study population**

In the discovery set, the mean age was 65.8 years (range, 35–87 years) for the cases, and 65.3 years (range, 34–89 years) for the controls. In the validation set, the mean age was 62.7 years (range, 32–95 years) for the cases, and 62.3 years (range, 33–92 years) for the controls (Supplementary Table 2). CRC patients had a higher percentage of a family history of cancer ($P < 0.001$ in the discovery set, $P = 0.039$ in the validation set) and a lower proportion of tea drinking ($P = 0.030$ in the validation set) than controls. There was no significant difference in other demographic characteristics (age, gender, BMI) or lifestyle factors (smoking, alcohol drinking) between cases and controls.

**Association between RP11-108K3.2 tagSNPs and the risk of colorectal cancer**

As shown in Table 1, six tagSNPs were included in the discovery set (rs2470151, rs3751591, rs3751592, rs77348531, rs17602308, and rs75106900), all of which conformed to the request of Hardy–Weinberg equilibrium. However, only the CT genotype of rs2470151 found a promising association after adjusting for age, gender, BMI, family history of cancer, smoking, alcohol consumption, and tea drinking (adjusted OR 0.74; 95% CI 0.51, 1.07; $P = 0.106$). The other tagSNPs were observed to have no meaningful association with the risk of CRC ($P > 0.20$).

The validation set carried on with the RP11-108K3.2 rs2470151 polymorphism, which was observed to be significantly associated with the risk of CRC in the validation set and the combined set (Table 2). In the validation set, subjects with the CT genotype had a significantly decreased risk of CRC compared with subjects with the CC genotype by a co-dominant model (adjusted OR 0.73; 95% CI 0.55, 0.97; $P = 0.030$), and this association was confirmed in the combined set (adjusted OR 0.73; 95% CI 0.58, 0.92; $P = 0.007$). However, the rs2470151 TT genotype did not show a significant difference, which may be due to the low genotype frequency. Also, we applied the dominant model to evaluate the effect of rs2470151 on the CRC risk; it still indicated the CT/TT genotype as being a protective factor for CRC in the combined set (adjusted OR 0.79; 95% CI 0.55, 0.97; $P = 0.025$). A subgroup analysis by gender is shown in Supplementary Table 3, which indicates that the association was stronger in females than males.

**Correlation between rs2470151 genotypes and RP11-108K3.2 expression**

Of the 71 CRC tissue samples, CC genotypes were identified in 18 cases (25.4%), CT genotypes in 40 cases (56.3%), and TT genotypes in 13 cases (18.3%). As shown in Figure 1, the relative expression level of RP11-108K3.2 was marginally higher in individuals carrying CC genotypes than in those carrying CT/TT genotypes (fold change $= 1.765$). The association between RP11-108K3.2 expression levels and clinic pathological characteristics was not statistically significant (Supplementary Table 4).

**Stratified analysis by selected variables between rs2470151 variant and the risk of CRC**

Stratified analyses by selected variables were conducted for the RP11-108K3.2 rs2470151 variant (Table 3). In the subjects without a family history of cancer, CT/TT carriers of rs2470151 had a lower risk of CRC compared with CC carriers (adjusted OR 0.77; 95% CI 0.62, 0.98; $P = 0.029$). In the combined set, smokers with the CT/TT genotypes...
Table 1. The association between tagSNPs of RP11-108K3.2 and the risk of colorectal cancer in the discovery set.

| Genotype | Cases, N (%) | Controls, N (%) | Crude OR (95% CI) | Adjusted<sup>a</sup> OR (95% CI) | P-value<sup>a</sup> | P-value<sup>b</sup> |
|----------|--------------|-----------------|-------------------|--------------------------|----------------|----------------|
| rs2470151 |              |                 |                   |                          |                 |                 |
| CC       | 121 (37.93)  | 103 (32.59)     | 1                 |                          |                 |                 |
| CT       | 143 (44.83)  | 155 (49.05)     | 0.88 (0.74, 1.06) | 0.173                    | 0.74 (0.51, 1.07) | 0.106 |
| TT       | 55 (17.24)   | 58 (18.35)      | 0.90 (0.71, 1.13) | 0.354                    | 0.80 (0.50, 1.29) | 0.367 |
| rs3751591 |              |                 |                   |                          |                 |                 |
| AA       | 243 (76.66)  | 231 (73.33)     | 1                 |                          |                 |                 |
| AG       | 67 (21.14)   | 75 (23.81)      | 0.92 (0.77, 1.11) | 0.393                    | 0.83 (0.56, 1.23) | 0.348 |
| GG       | 7 (2.21)     | 9 (2.86)        | 0.87 (0.56, 1.35) | 0.554                    | 0.66 (0.23, 1.87) | 0.430 |
| rs3751592 |              |                 |                   |                          |                 |                 |
| TT       | 239 (74.69)  | 239 (75.39)     | 1                 |                          |                 |                 |
| TC       | 74 (23.13)   | 74 (23.34)      | 1.00 (0.83, 1.20) | 1.000                    | 0.95 (0.65, 1.40) | 0.797 |
| CC       | 7 (2.19)     | 4 (1.26)        | 1.38 (0.63, 3.02) | 0.371                    | 1.76 (0.49, 6.30) | 0.383 |
| rs77348531|              |                 |                   |                          |                 |                 |
| AA       | 192 (60.00)  | 182 (57.05)     | 1                 |                          |                 |                 |
| AG       | 106 (33.13)  | 115 (36.05)     | 0.94 (0.79, 1.10) | 0.427                    | 0.91 (0.64, 1.29) | 0.581 |
| GG       | 22 (6.88)    | 22 (6.90)       | 0.97 (0.71, 1.33) | 0.867                    | 1.14 (0.59, 2.19) | 0.700 |
| rs17602308 |            |                 |                   |                          |                 |                 |
| TT       | 262 (82.13)  | 261 (82.86)     | 1                 |                          |                 |                 |
| TC       | 53 (16.61)   | 52 (16.51)      | 1.01 (0.82, 1.25) | 0.943                    | 1.04 (0.67, 1.62) | 0.846 |
| CC       | 4 (1.25)     | 2 (0.63)        | 1.50 (0.48, 4.66) | 0.420                    | 2.24 (0.40, 12.48) | 0.359 |
| rs75106900 |            |                 |                   |                          |                 |                 |
| C/C      | 189 (59.25)  | 178 (56.51)     | 1                 |                          |                 |                 |
| C/DEL    | 106 (33.23)  | 114 (36.19)     | 0.94 (0.79, 1.10) | 0.437                    | 0.89 (0.63, 1.26) | 0.518 |
| DEL/DEL  | 24 (7.52)    | 23 (7.30)       | 0.99 (0.73, 1.35) | 0.955                    | 1.19 (0.63, 2.24) | 0.598 |

BMI: body mass index; CI: confidence interval; OR: odds ratio.
*Adjusted for age, gender, BMI, occupation, education level, marital status, family history of cancer, smoking, alcohol consumption, and tea drinking.
*The tagSNPs with a P-value less than 0.20 in the discovery set were further analyzed in the validation set.
P < 0.20 is in bold, P < 0.05 is considered as statistically significant.

Table 2. Significant association between RP11-108K3.2 rs2470151 and the risk of colorectal cancer.

| Genotype | Cases, N (%) | Controls, N (%) | Crude OR (95% Cl) | Adjusted<sup>a</sup> OR (95% Cl) | P-value<sup>a</sup> | P-value<sup>b</sup> |
|----------|--------------|-----------------|-------------------|--------------------------|----------------|----------------|
| Discovery set |              |                 |                   |                          |                 |                 |
| CC       | 121 (37.93)  | 103 (32.59)     | 1                 |                          |                 |                 |
| CT       | 143 (44.83)  | 155 (49.05)     | 0.79 (0.56, 1.11) | 0.173                    | 0.74 (0.51, 1.07) | 0.106 |
| TT       | 55 (17.24)   | 58 (18.35)      | 0.81 (0.51, 1.27) | 0.354                    | 0.80 (0.50, 1.29) | 0.367 |
| CT + TT  | 198 (62.07)  | 213 (67.41)     | 0.79 (0.57, 1.10) | 0.159                    | 0.75 (0.53, 1.06) | 0.107 |
| Validation set |          |                 |                   |                          |                 |                 |
| CC       | 176 (53.48)  | 159 (29.94)     | 1                 |                          |                 |                 |
| CT       | 229 (46.17)  | 290 (54.61)     | 0.71 (0.54, 0.94) | 0.016                    | 0.73 (0.55, 0.97) | 0.030 |
| TT       | 91 (18.35)   | 82 (15.44)      | 1.00 (0.69, 1.45) | 0.989                    | 1.03 (0.71, 1.51) | 0.872 |
| CT + TT  | 320 (64.52)  | 372 (70.06)     | 0.78 (0.60, 1.01) | 0.058                    | 0.80 (0.61, 1.05) | 0.103 |
| Combined set |            |                 |                   |                          |                 |                 |
| CC       | 297 (36.44)  | 262 (30.93)     | 1                 |                          |                 |                 |
| CT       | 372 (45.64)  | 445 (52.54)     | 0.74 (0.59, 0.92) | 0.006                    | 0.73 (0.58, 0.92) | 0.007 |
| TT       | 146 (17.91)  | 140 (16.53)     | 0.92 (0.69, 1.22) | 0.566                    | 0.94 (0.70, 1.26) | 0.659 |
| CT + TT  | 518 (63.56)  | 585 (69.07)     | 0.78 (0.64, 0.96) | 0.017                    | 0.79 (0.64, 0.97) | 0.025 |

BMI: body mass index; CI: confidence interval; OR: odds ratio.
*Adjusted for age, gender, BMI, occupation, education level, marital status, family history of cancer, smoking, alcohol consumption, and tea drinking.
P < 0.20 was applied as the criteria to screen genotypes in the discovery set, and P < 0.05 was used as the criteria to testify rs2470151 genotype in the validation set.
P < 0.05 is in bold, P < 0.05 is considered as statistically significant.
had a lower risk of CRC than those with the CC genotype (adjusted OR 0.64; 95% CI 0.45, 0.92; \( P = 0.014 \)). Non-drinkers with the CT/TT genotypes had a significantly decreased risk of CRC compared with those with the CC genotype (adjusted OR 0.77; 95% CI 0.61, 0.99; \( P = 0.039 \)). In the non-tea drinkers, carrying the CT/TT genotypes seemed to be a protective factor for CRC (adjusted OR 0.70; 95% CI 0.53, 0.92; \( P = 0.011 \)).

**Discussion**

CRC is a major cause of morbidity and mortality worldwide.\(^{13} \) One of the fundamental processes driving the initiation and progression of CRC is the accumulation of a variety of genetic and epigenetic changes in the colorectum.\(^{8} \) Here, we explored the associations between the tag-SNPs on \( \text{RP11-108K3.2} \) and the risk of CRC through a two-stage, case-control study. This association between rs2470151 and the risk of CRC was influenced by selected variables including a family history of cancer, smoking, alcohol consumption, and tea drinking. Moreover, carrying the CC genotype of rs2470151 may increase \( \text{RP11-108K3.2} \) expression levels compared to the CT/TT carriers. To the best of our knowledge, this is the first study to explore the effect of \( \text{RP11-108K3.2} \) on the risk of CRC.

lncRNAs are implicated in serial steps of cancer development, interacting with DNA, RNA, protein molecules, and/or their combinations, acting as an essential regulator in chromatin organization, and in transcriptional and post-transcriptional regulation.\(^{14} \) There is increasing evidence revealing that SNPs on lncRNAs may be correlated with cancer, which may influence the expression of RNA or the stability of RNA conformation, resulting in the

---

**Table 3.** Risk of colorectal cancer associated with rs2470151 stratified by family history of cancer, smoking, alcohol consumption, and tea drinking.

| Variables                  | Genotype | Discovery set | Validation set | Combined set |
|----------------------------|----------|---------------|----------------|--------------|
|                            |          | \( \frac{N_{\text{cases}}}{N_{\text{controls}}} \) Adjusted OR (95% CI) | \( \frac{N_{\text{cases}}}{N_{\text{controls}}} \) Adjusted OR (95% CI) | \( \frac{N_{\text{cases}}}{N_{\text{controls}}} \) Adjusted OR (95% CI) |
| **Family history of cancer** |          |               |                |              |
| No                         | CC*      | 101/96        | 135/128        | 236/224      |
|                           | CT+TT    | 155/189 0.76 (0.53, 1.09) | 255/316 0.80 (0.59, 1.08) | 410/505 0.77 (0.62, 0.98) 0.029 |
| Yes                        | CC*      | 20/6          | 41/31          | 61/37        |
|                           | CT+TT    | 43/23 0.65 (0.21, 1.96) | 65/56 0.75 (0.40, 1.41) | 108/79 0.80 (0.47, 1.34) 0.389 |
| **Smoking**                |          |               |                |              |
| No                         | CC*      | 74/76         | 105/94         | 179/170      |
|                           | CT+TT    | 133/141 0.91 (0.60, 1.38) | 208/230 0.85 (0.60, 1.20) | 341/371 0.88 (0.67, 1.14) 0.324 |
| Yes                        | CC*      | 47/27         | 71/65          | 118/92       |
|                           | CT+TT    | 65/72 0.50 (0.27, 0.93) | 110/140 0.71 (0.46, 1.10) | 175/212 0.64 (0.45, 0.92) 0.014 |
| **Alcohol drinking**       |          |               |                |              |
| No                         | CC*      | 94/86         | 126/108        | 220/194      |
|                           | CT+TT    | 152/159 0.84 (0.57, 1.22) | 231/273 0.74 (0.54, 1.01) | 383/432 0.77 (0.61, 0.99) 0.039 |
| Yes                        | CC*      | 26/15         | 47/48          | 73/63        |
|                           | CT+TT    | 46/53 0.41 (0.17, 0.95) | 88/95 0.96 (0.57, 1.62) | 134/148 0.78 (0.51, 1.20) 0.259 |
| **Tea drinking**           |          |               |                |              |
| No                         | CC*      | 75/65         | 107/71         | 182/136      |
|                           | CT+TT    | 124/125 0.82 (0.53, 1.26) | 185/201 0.62 (0.43, 0.90) 0.011 | 309/326 0.70 (0.53, 0.92) 0.011 |
| Yes                        | CC*      | 46/38         | 68/88          | 114/126      |
|                           | CT+TT    | 74/88 0.63 (0.36, 1.11) | 130/165 1.13 (0.75, 1.69) | 204/253 0.92 (0.67, 1.28) 0.629 |

BMI: body mass index; CI: confidence interval; OR: odds ratio.

*Reference group.

\( P < 0.05 \) is in bold, \( P < 0.05 \) is considered as statistically significant.
modification of its interacting partners. SNP variants on lncRNAs were observed to function in different cancer types, including gastric cancer, lung cancer, prostate cancer, and breast cancer. Several studies also have reported that SNPs on lncRNAs are related to the risk of CRC. For example, Ling et al. reported that a novel lncRNA CCAAT2 was highly overexpressed in microsatellite-stable CRC, and the SNP status affected its expression, and the risk allele G of rs6983267 produced more CCAT2 transcript. Li et al. demonstrated that rs13252298 and rs1456315 on the lncRNA PRNCR1 were significantly associated with the decreased risk of CRC. Haiman et al. revealed that rs6983267 mapping on the 8q24 chromosomal region was also related to the risk of CRC. Based on the results of current studies, we hypothesized that the SNP rs2470151 may influence the expression of RP11-108K3.2, resulting in its subsequent biological processes. In our study, we observed that the CT/TT genotype carriers of rs2470151 polymorphism (C>T) had a significantly decreased risk of CRC compared with the subjects with CC genotype. Furthermore, different genotype carriers had different expression levels of RP11-108K3.2 in CRC tissues; however, the difference was not significant.

In previous epidemiological studies, there was quite a lot of evidence supporting the association between polymorphisms and different lifestyle factors, including smoking, alcohol, body fatness, diet, physical activity, and medications, in CRC. Bailie et al. also described several lifestyle factors associated with higher colorectal disease risk, most notably smoking and alcohol. Besides the factors above, family history is a well-established risk factor for CRC. In our study, we also demonstrated that the association between the risk of CRC and RP11-108K3.2 rs2470151 may be influenced by selected variables, including a family history of cancer, alcohol consumption, and tea drinking through stratified analyses.

There are several limitations in this study. First, although we observed a significant association between rs2470151 and the risk of CRC, how the genetic variability at this locus influences RP11-108K3.2 to function still remains to be fully elucidated. Second, the recall bias of lifestyle factors was inevitable in a case-control study. Third, most of the clinical information of the CRC cases was absent due to the population-based study. Fourth, the associations might have insufficient statistical power due to the relatively moderate sample size, and further studies with larger samples are expected to confirm our findings, especially in subgroup analysis and correlation study between rs2470151 genotypes and RP11-108K3.2 expression levels.

In summary, we have provided initial evidence that subjects carrying the rs2470151 CT/TT genotype had significantly lower CRC risk than those with the CC genotype. We also indicated that the CT/TT genotype of rs2470151 may have an effect on the lower expression level for RP11-108K3.2. This association between rs2470151 polymorphism and the risk of CRC may be influenced by a family history of cancer, alcohol consumption, and tea drinking. These results suggest that the SNP rs2470151 on RP11-108K3.2 may help to predict the CRC risk in Chinese populations.

Author contributions
Danjie Jiang and Mingjuan Jin contributed equally to this work.

Declaration of conflicting interest
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the National Natural Science Foundation of China (No.81673262 and 81373087), the National Key Basic Research Program (973 program No.2015CB554003), the Natural Science Foundation of Zhejiang Province, China (No. LQ15H1260001), and the Zhejiang Chinese Medical University Foundation (No.2019ZR03).

ORCID iD
Danjie Jiang https://orcid.org/0000-0002-7658-1548

Supplemental material
Supplemental material for this article is available online.

References
1. Brenner H, Kloor M and Pox CP. Colorectal cancer. Lancet 2014; 383: 1490–1502.
2. Al-Tassan NA, Whiffin N, Hosking FJ, et al. A new GWAS and meta-analysis with 1000Genomes imputation identifies novel risk variants for colorectal cancer. Sci Rep 2015; 5: 10442.
3. Peters U, Jiao S, Schumacher FR, et al. Identification of genetic susceptibility loci for colorectal tumors in a genomewide meta-analysis. Gastroenterology 2013; 144: 799–807.
4. Theodoratou E, Montazeri Z, Hawken S, et al. Systematic meta-analyses and field synopsis of genetic association studies in colorectal cancer. J Natl Cancer Inst 2012; 104: 1433–1457.
5. Xue Y, Gu D, Ma G, et al. Genetic variants in lncRNA HOTAIR are associated with risk of colorectal cancer. Mutagenesis 2015; 30: 303–310.
6. Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012; 22: 1775–1789.
7. Guttman M and Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature 2012; 482: 339–346.
8. Okugawa Y, Grady WM and Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology* 2015; 149: 1204–1225.

9. Zhang X, Zhou L, Fu G, et al. The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. *Carcinogenesis* 2014; 35: 2062–2067.

10. Slattery ML, Lundgreen A, Herrick JS, et al. Variation in the CYP19A1 gene and risk of colon and rectal cancer. *Cancer Causes Control* 2011; 22: 955–963.

11. Lin JH, Manson JE, Kraft P, et al. Estrogen and progesterone-related gene variants and colorectal cancer risk in women. *BMC Med Genet* 2011; 12: 78.

12. Zhang Y, Liu B, Jin M, et al. Genetic polymorphisms of transforming growth factor-beta1 and its receptors and colorectal cancer susceptibility: a population-based case-control study in China. *Cancer Lett* 2009; 275: 102–108.

13. Dickinson BT, Kisiel J, Ahlquist DA, et al. Molecular markers for colorectal cancer screening. *Gut* 2015; 64: 1485–1494.

14. Han D, Wang M, Ma N, et al. Long noncoding RNAs: novel players in colorectal cancer. *Cancer Lett* 2015; 361: 13–21.

15. Du M, Wang W, Jin H, et al. The association analysis of lncRNA HOTAIR genetic variants and gastric cancer risk in a Chinese population. *Oncotarget* 2015; 6: 31255–31262.

16. Pan W, Liu L, Wei J, et al. A functional lncRNA HOTAIR genetic variant contributes to gastric cancer susceptibility. *Mol Carcinog* 2016; 55: 90–96.

17. Hu L, Chen SH, Lv QL, et al. Clinical significance of long non-coding RNA CASC8 rs10505477 polymorphism in lung cancer susceptibility, platinum-based chemotherapy response, and toxicity. *Int J Environ Res Public Health* 2016; 13.

18. Yuan H, Liu H, Liu Z, et al. A novel genetic variant in long non-coding RNA gene NEXN-AS1 is associated with risk of lung cancer. *Sci Rep* 2016; 6: 34234.

19. Guo H, Ahmed M, Zhang F, et al. Modulation of long non-coding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. *Nat Genet* 2016; 48: 1142–1150.

20. Jin G, Sun J, Isaacs SD, et al. Human polymorphisms at long non-coding RNAs (lncRNAs) and association with prostate cancer risk. *Carcinogenesis* 2011; 32: 1655–1659.

21. Ingle JN, Xie F, Ellis MJ, et al. Genetic polymorphisms in the long noncoding RNA MIR2052HG offer a pharmacogenomic basis for the response of breast cancer patients to aromatase inhibitor therapy. *Cancer Res* 2016; 76: 7012–7023.

22. Yan R, Cao J, Song C, et al. Polymorphisms in lncRNA HOTAIR and susceptibility to breast cancer in a Chinese population. *Cancer Epidemiol* 2015; 39: 978–985.

23. Ling H, Spizzo R, Atlasi Y, et al. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. *Genome Res* 2013; 23: 1446–1461.

24. Li L, Sun R, Liang Y, et al. Association between polymorphisms in long non-coding RNA PRNCR1 in 8q24 and risk of colorectal cancer. *J Exp Clin Cancer Res* 2013; 32: 104.

25. Haiman CA, Le Marchand L, Yamamato J, et al. A common genetic risk factor for colorectal and prostate cancer. *Nat Genet* 2007; 39: 954–956.

26. Andersen V, Egeberg R, Tjonneland A, et al. Interaction between interleukin-10 (IL-10) polymorphisms and dietary fibre in relation to risk of colorectal cancer in a Danish case-cohort study. *BMC Cancer* 2012; 12: 183.

27. Andersen V, Holst R, Kopp TI, et al. Interactions between diet, lifestyle and IL10, IL1B, and PTGS2/COX-2 gene polymorphisms in relation to risk of colorectal cancer in a prospective Danish case-cohort study. *PLoS One* 2013; 8: e78366.

28. Kopp TI, Andersen V, Tjonneland A, et al. Polymorphisms in ATP-binding cassette transporter genes and interaction with diet and lifestyle factors in relation to colorectal cancer in a Danish prospective case-cohort study. *Scand J Gastroenterol* 2015; 50: 1469–1481.

29. Kopp TI, Andersen V, Tjonneland A, et al. Polymorphisms in NFKB1 and TLR4 and interaction with dietary and lifestyle factors in relation to colorectal cancer in a Danish prospective case-cohort study. *PLoS One* 2015; 10: e0116394.

30. Baille L, Loughrey MB and Coleman HG. Lifestyle Risk Factors for Serrated Colorectal Polyps: A Systematic Review and Meta-analysis. *Gastroenterology* 2017; 152: 92–104.

31. Fuchs CS, Giovannucci EL, Colditz GA, et al. A prospective study of family history and the risk of colorectal cancer. *N Engl J Med* 1994; 331: 1669–1674.