Case Control Study

Association of twelve polymorphisms in three onco-lncRNA genes with hepatocellular cancer risk and prognosis: A case-control study

Ben-Gang Wang, Qian Xu, Zhi Lv, Xin-Xin Fang, Han-Xi Ding, Jing Wen, Yuan Yuan

Ben-Gang Wang, Qian Xu, Zhi Lv, Xin-Xin Fang, Han-Xi Ding, Jing Wen, Yuan Yuan, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, the First Hospital of China Medical University, and Key Laboratory of Cancer Etiology and Prevention (China Medical University), Liaoning Provincial Education Department, Shenyang 110001, Liaoning Province, China

Ben-Gang Wang, Hepatobiliary Surgery Department of General Surgery Institute, the First Hospital of China Medical University, Shenyang 110001, Liaoning Province, China

ORCID number: Ben-Gang Wang (0000-0002-5630-1748); Qian Xu (0000-0002-3082-7530); Zhi Lv (0000-0001-8443-169X); Xin-Xin Fang (0000-0002-7617-5099); Han-Xi Ding (0000-0002-8208-541X); Jing Wen (0000-0002-0629-4631); Yuan Yuan (0000-0002-7394-9036).

Author contributions: Yuan Y conceived and designed this study; Wang BG, Lv Z, Fang XX, and Wen J preformed the experiments; Wang BG, Xu Q, and DingHX were responsible for the data analysis and performed data interpretation; Wang BG wrote the paper; Lv Z and Yuan Y revised the manuscript.

Supported by the Natural Science Foundation of Liaoning Province in China, No. 20170541001.

Institutional review board statement: The study was approved by the ethics committee of the First Hospital of China Medical University.

Informed consent statement: All patients gave informed consent.

Conflict-of-interest statement: All authors disclose no conflicts of interest that may bias their work.

STROBE statement: The guidelines of the STROBE Statement have been adopted.

Open-Access: This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

Manuscript source: Unsolicited manuscript

Correspondence to: Yuan Yuan, MD, PhD, Professor, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Hospital of China Medical University, No.155 NanjingBei Street, Heping District, Shenyang 110001, Liaoning Province, China. yuan@cmu.edu.cn
Telephone: +86-24-83282153
Fax: +86-24-83282383

Received: March 22, 2018
Peer-review started: March 23, 2018
First decision: April 10, 2018
Revised: April 17, 2018
Accepted: May 18, 2018
Article in press: May 18, 2018
Published online: June 21, 2018

Abstract

AIM
To evaluate the association of 12 tag single nucleotide polymorphisms (tagSNPs) in three onco-long non-coding RNA (lncRNA) genes (HOTTIP, CCAT2, MALAT1) with the risk and prognosis of hepatocellular cancer (HCC).

METHODS
Twelve tagSNPs covering the three onco-lncRNAs were genotyped by the KASP method in a total of 1338 samples, including 521 HCC patients and frequency-matched 817 controls. The samples were obtained from an unrelated Chinese population at the First Hospital of
RESULTS
Three SNPs in HOTTIP, one promoter SNP in MALAT1, and one haplotype of HOTTIP were associated with HCC risk. The HOTTIP rs17501292, rs2067087, and rs17427960 SNPs were increased to 1.55-, 1.20-, and 1.18-fold HCC risk under allelic models ($P = 0.012, 0.017$ and 0.049, respectively). MALAT1 rs4102217 SNP was increased to a 1.32-fold HCC risk under dominant models ($P = 0.028$). In addition, the two-way interaction of HOTTIP rs17501292-MALAT1 rs619586 polymorphisms showed a decreased effect on HCC risk ($P_{\text{interaction}} = 0.028, OR = 0.30$) and epistasis with each other. HOTTIP rs3807598 variant genotype showed significantly longer survival time in HBV negative subgroup ($P = 0.049, HR = 0.12$), and MALAT1 rs591291 showed significantly better prognosis in female and HBV negative subgroups ($P = 0.022, HR = 0.37; P = 0.042, HR = 0.25$, respectively).

In the study, no significant effect was observed in eQTL analysis.

CONCLUSION
Specific IncRNA (HOTTIP and MALAT1) SNPs have potential to be biomarkers for HCC risk and prognosis.

Key words: Hepatocellular cancer; Single nucleotide polymorphism; Long non-coding RNA; Risk; Prognosis

© The Author(s) 2018. Published by Baishideng Publishing Group Inc. All rights reserved.

INTRODUCTION
Hepatocellular cancer (HCC) is a common malignant tumor with high incidence and mortality, and it is the main histologic type of primary liver cancer[1,2]. Similar to most other solid tumors, HCC patients are considered to be incurable due to the extensive heterogeneity in the clinical manifestations and biological characteristics[3]. HCC heterogeneity can be manifested by diverse genetic, epigenetic, and histogenic features, as well as ethnic differences in patients[4]. Knowledge of genetic and epigenetic variations could aid the early detection and personalized management of HCC. To date, the research hotspots regarding genetic and epigenetic variations are not only coding genes but also noncoding RNAs.

Long non-coding RNAs (IncRNAs) are a type of noncoding RNA with a length of 200 bp, which can function as miRNA sponges to compete with mRNAs by acting as so-called competing endogenous RNAs (ceRNAs)[5]. Genetic variations such as single-nucleotide polymorphisms (SNPs) can alter the expression of coding genes and IncRNAs[6]. To date, several IncRNAs have been reported to be involved in carcinogenesis, such as H19, HOXATIR, HOTTIP, CCAT2, and MALAT1. Regarding studies of genetic variation, only H19 and HOXATIR SNPs have been well investigated. For example, several meta-analyses showed that H19 and HOXATIR SNPs were associated with cancer risk[7-9], and the HOXATIR rs920778 SNP was found to be associated with ovarian cancer prognosis[10]. Only four studies have focused on the polymorphisms of the onco-IncRNAs HOTTIP, CCAT2, and MALAT1 (Gene ID: 100316868, 101805488, 378938, respectively)[11-14]. Among them, Gong et al[11] found that these SNPs were significantly associated with lung cancer susceptibility or platinum-based chemotherapy response. However, no SNPs in the above-mentioned IncRNAs have been reported to be associated with HCC risk, and few comprehensive and systematic analyses have been performed on polymorphisms in these three onco-IncRNAs. It thus remains unclear whether the promising SNPs in these IncRNAs have potential to be used as biomarkers for HCC risk and prognosis.

In the present study, we adopted a candidate gene association study strategy with the selected 12 potentially functional tagSNPs covering the three onco-IncRNAs HOTTIP, CCAT2, and MALAT1 to determine whether these SNPs are associated with HCC risk and prognosis and whether promising SNPs could affect the expression of corresponding IncRNAs. We aimed to identify predictive biomarkers for HCC risk and prognosis, establish experimental basis for the comprehension of the HCC etiology, and improve our understanding of the pathogenesis and disease progression of HCC.

MATERIALS AND METHODS
Patients and study design
This research project was approved by the Ethical Committee of the First Hospital of China Medical University and written informed consent was obtained from each subject at the time of recruitment. The study

Wang BG, Xu Q, Lv Z, Fang XX, Ding HX, Wen J, Yuan Y. Association of twelve polymorphisms in three onco-lncRNA genes with hepatocellular cancer risk and prognosis: A case-control study. World J Gastroenterol 2018; 24(23): 2482-2490 Available from: URL: http://www.wjgnet.com/1007-9327/full/v24/i23/2482.htm DOI: http://dx.doi.org/10.3748/wjg.v24.i23.2482
was designed to be composed of two parts: risk and prognosis. In the risk study, a total of 1338 participants were recruited, including 521 patients who underwent surgical operation for HCC at the First Hospital of China Medical University between 2012 and 2015. The inclusion and exclusion criteria were as follows: (1) The participants who underwent surgical operation were diagnosed with HCC by pathological confirmation, in accordance with the World Health Organization (WHO) classification; and (2) removal of other pathological types of liver cancer (gallbladder cell carcinoma, mix-type liver cancer, and hepatosarcoma). A total of 817 frequency-matched controls were also recruited, some of whom were from a health screening program from the Zhuanghe area, Liaoning Province, China, performed between 2002 and 2012, while others were from a health screening program at the First Hospital of China Medical University performed between 2012 and 2015.

To investigate further the association of these IncRNA polymorphisms with clinicopathological parameters and overall survival of HCC patients, we used data of 351 HCC cases for which information on death or survival was available. Patients with (1) distant metastasis found preoperatively or (2) incomplete pathological data entries were excluded from the survival analysis. Follow-up was completed by July 1st, 2017. For the promising SNPs, IncRNA expression was investigated to explore the possible mechanism by which they exerted their effects, according to our experimental and bioinformatic data. The study design is shown in Figure 1. This research project was approved by the Ethical Committee of the First Hospital of China Medical University and written informed consent was obtained when the patients and controls were recruited.

Selected polymorphic sites
We selected polymorphisms using 1000G data (http://www.internationalgenome.org/home), as reported previously[15-17]. The tagSNPs were selected separately using the following criteria: (1) Haploview with the Tagger function was used; (2) the population of the HapMap selected Chinese Han Beijing (CHB) population; (3) those for which pairwise tagging had $r^2$ of $\geq 0.8$; and (4) those with a minor allele frequency of $\geq 5\%$. The selection area was enlarged by 10 kb both upstream and downstream for these three IncRNA genes. FastSNP and fSNP searches were used to predict the potential SNP function (http://compbio.cs.queensu.ca/F-SNP/)[18,19]. A total of 12 SNPs from the three IncRNA genes were selected by integrating these two publicly available tools. Locations and characteristics of the selected SNPs are shown in Table 1.

Genotyping
Genomic DNA was extracted using a previously reported method[20] and diluted to a working concentration of 20 ng/µL for genotyping. The genotyping assay was performed by Gene Company (Shanghai, China), using allele-specific PCR with KASPar (KASP) reagents (LGCGenomics, Hoddesdon, United Kingdom). For quality control, we repeatedly genotyped 10% of the total samples at one time. The concordance rate of these repeated samples reached 100%, which demonstrated that the genotyping results were reliable.

eQTL analyses
The extraction of total RNA from 68 HCC specimens and corresponding samples from nearby noncancerous regions was performed as described previously[15], and a total of 2.0 µg of isolated RNA was converted into cDNA using Quantscript RT Kit (Tiangen Biotech, Beijing, China). The RNA expression levels of the promising IncRNA genes (HOTTIP and MALAT1) and an internal-control gene (GAPDH) were measured using SYBR Premix EX Taq II (TaKaRa Biotech, Dalian, China) in an Eppendorf Mastercycler Gradient System (Eppendorf AG, Hamburg, Germany). Each reaction was performed in duplicate and controls without a template were also tested every time. The primers are summarized in Supplementary Table 1.

To perform functional candidate polymorphism and expression quantitative trait locus (eQTL) analyses on the promising genes, we mined the data from the following databases: GTExPortal (https://www.gtexportal.org/home/) and HaploReg (http://www.broadinstitute.org/mammals/haploreg/haploreg.php)[21].

Statistical analysis
Between-group differences in sex variability, as well as accordance with Hardy-Weinberg equilibrium, were
Table 1 Association of lncRNA gene single nucleotide polymorphisms and risk of hepatocellular cancer (n)

| Gene  | Chr. Pos. | SNP1 | Loc. | Genotype | Controls (%) | Cases (%) | P2 value | OR (95%CI) | Pmeta value |
|-------|-----------|------|------|----------|--------------|-----------|----------|------------|-------------|
| HOTTIP| 7p15.2    | rs17501292 | Exon | TT       | 732 (91.2) | 453 (87.1) | 0.021    | 1.52 (1.06-2.17) | 0.190       |
|       |           |       |      | TG       | 71 (8.8)   | 66 (12.7)  | NA       | NA         |             |
|       |           |       |      | GC       | 0 (0)      | 1 (0.2)    | NA       | NA         |             |
|       | rs2067087 | Exon |      | GG       | 174 (21.7) | 88 (16.9)  | 0.012    | 1.53 (1.10-2.18) | 0.674       |
|       |           |       |      | GC       | 405 (50.6) | 263 (50.7) | 0.236    | 1.16 (0.90-1.49) |             |
|       |           |       |      | CC       | 222 (27.7) | 168 (32.4) | 0.015    | 1.49 (1.08-2.08) |             |
|       | rs17427960| Intron|      | CC+GC    | 172 (21.7) | 85 (16.7)  | 0.035    | 1.35 (1.02-1.82) | 0.613       |
|       |           |       |      | C+G      | 387 (48.8) | 259 (50.9) | 0.707    | 1.03 (0.81-1.35) |             |
|       |           |       |      | AA       | 254 (32.5) | 165 (32.4) | 0.032    | 1.78 (1.03-2.00) |             |
| MALAT1| 22q13.2   | rs4102217 | Promoter | GG       | 608 (75.1) | 362 (69.6) | 0.049    | 1.18 (1.00-1.37) | 0.055       |
|       |           |       |      | GC       | 180 (22.2) | 148 (28.5) | 0.011    | 1.39 (1.08-1.79) |             |
|       |           |       |      | CC       | 22 (2.7)   | 10 (1.9)   | 0.481    | 0.76 (0.36-1.63) |             |
|       |           |       |      | GC+CC    | 0.028      | 0.028      | 0.028    | 1.32 (1.03-1.69) |             |

1The sort order was according to the SNP location in its genes from 5' starting to 3' ends. 2P value was calculated by adjusted by age and gender. NA: Not available; Chr. Pos.: Chromosomal position; Loc.: Localisation; Pmeta: P value for Hardy-Weinberg Equilibrium.

compared by the χ2 test and by analysis of variance for age variability. Multivariate logistic regression with adjustments for age and sex was used to show the association between selected lncRNA polymorphisms and HCC risk. The haplotypes of each gene were analyzed using SHEsis software[22]. The two-way pairwise interactions of lncRNA SNP-SNP were calculated using multivariate logistic regression. Univariate and multivariate survival analyses were carried out by the log-rank test and the Cox proportional hazards model. The differences of relative lncRNA levels between two groups were tested by Student's t-test. P value < 0.05 was considered to be significant.

RESULTS

The association of SNPs in lncRNA genes with HCC risk

The demographic characteristics of HCC cases and control subjects are shown in Supplementary Table 2. All polymorphism genotype distributions in cases and controls are shown in Table 1, including 12 SNPs in three lncRNA genes (HOTTIP: rs3807598, rs17501292, rs2067087, rs17427960, rs78248039; CCAT2: rs3843549, rs138947056, rs6983267; MALAT1: rs4102217, rs591291, rs11227209, and rs619586). Among them, most SNPs accorded with Hardy-Weinberg equilibrium, except for CCAT2 rs6983267 (P_HWE = 0.029). This SNP was thus excluded from further association analysis.

Among these remaining 11 SNPs, four in the HOTTIP and MALAT1 genes were associated with HCC risk, and they all increased HCC risk (HOTTIP: rs17501292, rs2067087, rs17427960; MALAT1: rs4102217; Table 1, Supplementary Table 3). Among them, two SNPs (HOTTIP: rs17501292; MALAT1: rs4102217) showed significant in a dominant model and the other two showed in a recessive model. HOTTIP rs17501292 and MALAT1 rs4102217 were associated with an increased risk of HCC (P = 0.017 and 0.028, OR = 1.54 and 1.32, respectively) in a dominant model. In addition, HOTTIP rs2067087 and rs17427960 variant genotypes also showed associations with an increased risk of HCC (P = 0.035 and 0.028, OR = 1.35 and 1.39, respectively) in a recessive model. Stratified analysis based on gender, age, smoking, and drinking was performed to analyze the association between each SNP and HCC risk. The results are shown in Supplementary Table 4 and suggest these variables have potential predictive value for specific subgroup populations in HCC risk.

The association of haplotype in four lncRNA genes with HCC risk

We chose to exclude haplotypes with a frequency of less than 0.03 from the analysis. We found only one haplotype in the HOTTIP gene that was associated with HCC risk. Compared with other haplotypes, patients with the C-G-T-A haplotype of HOTTIP rs3807598-rs17501292-rs2067087-rs17427960 showed a 1.91-fold increased risk of HCC (P = 0.006, 95%CI: 1.20-3.05; Table 2).

Two-way SNP-SNP interaction models for lncRNA polymorphisms

For the data mining of two-way SNP-SNP interactions, we analyzed all possible pair combinations between all of these 11 SNPs and found that the pairwise interaction of HOTTIP rs17501292-MALAT1 rs619586 was significant (P_interaction = 0.028, OR = 0.30, 95%CI: 0.10-0.88; Table 3).

We further analyzed the epistatic effect of HOTTIP rs17501292 and MALAT1 rs619586 and found in the
subset with MALAT1 rs619586 AA wild type and HOTTIP rs17501292 SNP an increased risk of HCC under a dominant model (\(P = 0.002, \text{OR} = 1.85\)); however, in the subset with the HOTTIP rs17501292 TG+GG genotype, the MALAT1 rs619586 SNP decreased the risk of HCC under a dominant model (\(P = 0.050, \text{OR} = 0.36\); Supplementary Table 5).

The association of lncRNA SNPs with HCC cancer prognosis

We analyzed the association of all 11 SNPs with the overall survival of HCC patients but found no significant association in either univariate or multivariate Cox proportional hazard analysis (Supplementary Table 6). In the stratified analysis, those with the HOTTIP rs3807598 variant genotype were shown to have a significantly longer survival time in the HBV-negative subgroup (\(P = 0.049, \text{HR} = 0.12, 95\% \text{CI} = 0.02-0.99\)), and MALAT1 rs591291 showed an association with a significantly better prognosis in the female and HBV-negative subgroups (\(P = 0.022, \text{HR} = 0.37, 95\% \text{CI} = 0.16-0.87\); \(P = 0.042, \text{HR} = 0.25, 95\% \text{CI} = 0.07-0.95\), respectively; Table 4, Supplementary Table 7).

DISCUSSION

In this study, we preliminarily screened all of the tagSNPs covering three onco-lncRNAs, HOTTIP, CCAT2, and MALAT1, for associations with HCC risk and prognosis. We identified four promising risk-associated SNPs, one haplotype, and a two-way pairwise interaction combination associated with HCC risk. We also found that patients carrying the HOTTIP rs3807598 and MALAT1 rs591291 variant genotypes had a longer survival time in the HBV-negative subgroup. Further molecular experiments were also conducted to investigate whether the tagSNPs could affect the expression of the corresponding lncRNAs. Our study provides an experimental basis for seeking predictive biomarkers for the risk and prognosis of HCC.

| Variables | MALAT1 rs619586 |
|-----------|-----------------|
| HOTTIP rs17501292 TT | AA | AG + GG |
| Case/control | 372/617 | 78/107 |
| OR (95%CI) | 1 | 1.21(0.88-1.66) |
| TG + GG Case/control | 61/55 | 6/15 |
| OR (95%CI) | 1.84 (1.25-2.71) | 0.66 (0.26-1.73) |
| \(P_{\text{interaction}} = 0.028, \text{OR (95\%CI) = 0.30 (0.10-0.88)} \) |

HCC: Hepatocellular cancer; CON: Control.

eQTL analysis

We used eQTL analysis to investigate the effect of the SNPs identified to be associated with HCC risk on lncRNA expression. In neither the cancerous group nor the noncancerous group was a significant difference observed for the effect of the positive SNPs on lncRNA expression levels (Table 5). Among them, only the heterozygote genotype of intronic rs17427960 of the HOTTIP gene was associated with higher lncRNA-HOTTIP expression, with borderline significance (CA vs CC: \(P = 0.063; \text{Table 5})

| Variables | Malat1 rs619586 |
|-----------|---------------|
| Haplotype for 1 HOTTIP rs3807598-rs17501292-rs2067087-rs17427960; 2 CCAT2 rs3843549-rs138947056-rs6983267; 3 MALAT1 rs4102217-rs591291-rs11227209-rs619586.

Haplotype for 1 HOTTIP rs3807598-rs17501292-rs2067087-rs17427960; 2 CCAT2 rs3843549-rs138947056-rs6983267; 3 MALAT1 rs4102217-rs591291-rs11227209-rs619586.

### Table 2: Association of haplotype of lncRNA gene and hepatocellular cancer risk (n)

| Haplotype | Control (%) | Case (%) | OR (95% CI) | \(P\) value |
|-----------|-------------|----------|-------------|------------|
| HOTTIP    |             |          |             |            |
| CGCA      | 33.54 (2.4) | 407.00 (4.5) | 1.91 (1.20-3.05) | 0.006 |
| CTGC      | 663.39 (47.3) | 389.97 (43.4) | 0.85 (0.71-1.01) | 0.066 |
| GTCA      | 604.75 (43.1) | 401.77 (44.7) | 1.08 (0.91-1.28) | 0.406 |
| CCAT2     |             |          |             |            |
| AAG       | 623.60 (39.8) | 434.27 (42.7) | 1.13 (0.96-1.33) | 0.136 |
| AAT       | 937.40 (59.8) | 577.73 (56.9) | 0.89 (0.75-1.04) | 0.136 |
| MALAT1    |             |          |             |            |
| CTCA      | 217.99 (13.8) | 167.00 (16.3) | 1.22 (0.08-1.52) | 0.080 |
| GCCA      | 934.81 (59.3) | 599.74 (58.6) | 0.97 (0.82-1.14) | 0.718 |
| GTCA      | 296.19 (18.8) | 1720 (16.6) | 0.86 (0.70-1.06) | 0.160 |
| GTG       | 85.00 (5.4) | 58.94 (5.8) | 1.07 (0.76-1.51) | 0.689 |

Haplotype for 1 HOTTIP rs3807598-rs17501292-rs2067087-rs17427960; 2 CCAT2 rs3843549-rs138947056-rs6983267; 3 MALAT1 rs4102217-rs591291-rs11227209-rs619586.

| Variables | MALAT1 rs619586 |
|-----------|---------------|
| HOTTIP rs17501292 TT | AA | AG + GG |
| Case/control | 372/617 | 78/107 |
| OR (95%CI) | 1 | 1.21(0.88-1.66) |
| TG + GG Case/control | 61/55 | 6/15 |
| OR (95%CI) | 1.84 (1.25-2.71) | 0.66 (0.26-1.73) |
| \(P_{\text{interaction}} = 0.028, \text{OR (95\%CI) = 0.30 (0.10-0.88)} \) |

HCC: Hepatocellular cancer; CON: Control.
LncRNAs function as ceRNAs to compete with mRNAs for access to miRNAs, which could regulate the expression of coding genes. Most studies on lncRNA expression have focused on H19 and HOTAIR as well. They can also promote HCC metastasis and epithelial-mesenchymal transition. The HOTTIP gene is located in 7p15.2 and has three exons, the CCAT2 gene is located in 8q24.21 and has one exon, and the MALAT1 gene is located in 11q13.1 and has two exons. The most common SNPs reported for these genes are HOTTIP rs3807598, CCAT2 rs6983267, and MALAT1 rs619586. The first of these was found to be predictive of hematological toxicity in a three-way interaction pattern, and the latter two were indicated to be associated with platinum-based...
Wang BG et al. IncRNA SNPs with HCC risk and prognosis

chemotherapy response in lung cancer[11]. In this study, we found that SNPs in two exons (rs17501292 and rs2067087) and one intron (rs17427960) of the HOTTIP gene as well as an SNP in the promoter (rs4102217) of the MALAT1 gene were associated with HCC risk; these variant alleles increased HCC risk in the range from 1.18- to 1.55-fold. These four SNPs are reported here for the first time to be associated with cancer risk. Concerning the commonly studied HOTTIP rs3807598 and MALAT1 rs619586 SNPs, no significant associations with HCC risk were found in this study, which is consistent with the findings in a report by Liu[35]. In addition, we found that none of the CCAAT2 SNPs was associated with HCC risk. Following the identification of the possible significant SNPs, we further analyzed the relationship between the HOTTIP C-G-C-A haplotype of rs3807598-rs17501292-rs2067087-rs17427960 and HCC risk. The results showed an increase in HCC risk of 1.91-fold in those with this haplotype, and the OR value was greater than that for each SNP alone. Taking these findings together, it is newly indicated that the HOTTIP SNPs rs17501292, rs2067087, and rs17427960 and the MALAT1 SNP rs4102217 have potential to be biomarkers for HCC risk.

Combined interaction analysis for multiple SNPs from different genes is more sensitive and powerful than one-dimensional SNP analysis[36]. For individual SNPs at single loci that were previously shown to have no or a weak effect on disease risk, an epistatic effect may appear when they are analyzed in combination[27]. One of the most significant findings in this study was the SNP-SNP interaction identified for the HOTTIP rs17501292-MALAT1 rs619586 polymorphisms, which was confirmed by the epistatic effect analysis. In the main-effect analysis, HOTTIP rs17501292 had a weak effect, and MALAT1 rs619586 had no effect on the risk of HCC. However, the pairwise analysis of these two in combination showed that they had an interactive effect on HCC risk. Subsequently, we analyzed the epistatic effect of these two SNPs and found that MALAT1 rs619586 was associated with a decreased risk of HCC only in the presence of the HOTTIP rs17501292 TG+GG genotype. A similar epistatic effect between coding genes was also found in our previous study[38]. Further investigations are needed to verify our findings and the mechanism involved in the epistatic phenomenon.

In the prognostic analysis, we found no significant association of the studied SNPs with the overall survival of HCC patients. However, in the stratified analysis based on gender, we found that MALAT1 rs591291 was associated with significantly better prognosis in the female subgroup. When stratified by HBV infection status, we found that patients carrying HOTTIP rs3807598 and MALAT1 rs591291 variant genotypes had longer survival times. As some biomarkers are specific for certain subgroups and have potential to be used for the diagnosis or individualized therapy of specific subgroups[39], the above-mentioned polymorphisms could have value in predicting HCC prognosis for certain subgroups.

eQTL is an analysis in which the combination of mRNA expression and genotype data is applied to determine which variants are correlated with the transcription levels of genes[40]. We analyzed the SNPs potentially associated with HCC risk in our own data and then reanalyzed them in two public databases for the eQTL analysis. HOTTIP rs17501292 and rs2067087 are both located in exon 2 of this gene, while rs17427960 is in intron 2. In contrast, rs4102217 is located at -1255 bp of the MALAT1 gene, within the promoter region. Among these four SNPs, we found that only the heterozygous genotype of intronic rs17427960 of the HOTTIP gene was associated with a higher IncRNA-HOTTIP expression level, with borderline significance. The public databases offered some supportive evidence for this from findings in other tissues, suggesting that rs4102217 in the MALAT1 promoter is a functional SNP in 34 different tissues, such as pancreas and stomach, and that exonic rs17501292 of HOTTIP is a functional SNP in tibial artery tissue. In addition, some regulatory motifs, which were predicted by the bioinformatical software listed in Supplementary Figure 1, are transcription factors like PAX-4 and AP1. Thus, it is reasonable to assume that these SNPs could regulate certain motifs, leading to higher expression of oncogenic IncRNA and thus an elevation of HCC risk. However, further functional research is required to confirm this.

In summary, we found that the SNPs rs17501292, rs2067087, and rs17427960 in the HOTTIP gene, rs4102217 in the MALAT1 gene, and a haplotype of HOTTIP increased the risk of HCC. In addition the SNPs HOTTIP rs3807598 and MALAT1 rs591291 were associated with longer survival time in the HBV-negative subgroup.

ARTICLE HIGHLIGHTS

Research background
Genetic polymorphisms could be biomarkers for cancer risk and prognosis. Recent years, it was found that coding genes and non-coding genes had single nucleotide polymorphisms (SNPs). lncRNAs had important roles in the tumor incidence, progression, and prognosis. Thus, lncRNA polymorphisms had potential to be biomarkers for cancer precaution and prognostic prediction.

Research motivation
The aim of this study is to screen out the effective biomarkers for the hepatocellular cancer (HCC) risk and prognosis. The selected polymorphisms would have potential for the prediction of cancer risk and prognosis.

Research objectives
Five hundred and twenty-one patients of HCC and frequency matched 817 controls were studied for the cancer risk study. Among them, 351 patients for which the information was all available were recruited for the prognosis study. Then, 68 HCC specimens and corresponding samples from the noncancerous region were detected for the expression level study.

Research methods
For the risk and prognosis study, the samples were detected by the genomic DNA extracted and allele-specific PCR with KASPar reagents. The single nucleotide polymorphisms were selected by the Haploview software. The expression level study isolated the RNA isolated and then converted it to cDNA.
REFERENCES

1. Bosetti C, Turati F, La Vecchia C. Hepatocellular carcinoma epidemiology. *Best Pract Res Clin Gastroenterol* 2014; 28: 575-577 [PMID: 25266006 DOI: 10.1016/j.bpg.2014.08.007].

2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008: Int J Cancer 2010; 127: 2893-2917 [PMID: 21351269 DOI: 10.1002/jic.25516].

3. Roessler S, Budhu A, Wang XW. Deciphering cancer heterogeneity: the biological space. *Front Cell Dev Biol* 2014; 2: 12 [PMID: 25636720 DOI: 10.3389/fcell.2014.00012].

4. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell* 2013; 153: 38-55 [PMID: 23540689 DOI: 10.1016/j.cell.2013.03.008].

5. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, Chen N, Sun F, Fan Q. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res* 2010; 38: 5366-5383 [PMID: 20423907 DOI: 10.1093/nar/gkq285].

6. Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, Zeng Y, Miao R, Jin G, Ma H, Chen Y, Shen H. Genetic variants of mRNA sequences and non-small cell lung cancer survival. *J Clin Invest* 2008; 118: 2600-2608 [PMID: 18521189 DOI: 10.1172/JCI34934].

7. Li XF, Yin XH, Cai JW, Wang MJ, Zeng YQ, Li M, Niu YM, Shen M. Significant association between IncRNA H19 polymorphisms and cancer susceptibility: a meta-analysis. *Onco Targets Ther* 2017; 8: 4513-4513 [PMID: 28404885 DOI: 10.18632/oncotarget.16658].

8. Lv Z, Xu Q, Yuan Y. A systematic review and meta-analysis of the association between long non-coding RNA polymorphisms and cancer risk. *Mutat Res* 2017; 771: 1-14 [PMID: 28342449 DOI: 10.1016/j.mrrev.2016.10.002].

9. Chu H, Chen Y, Yuan Q, Hua Q, Zhang X, Wang M, Tong N, Zhang W, Chen J, Zhang Z. The HOTAIR, PRRCR1 and P0L2RE polymorphisms are associated with cancer risk: a meta-analysis. *Onco Targets Ther* 2017; 8: 43271-43283 [PMID: 28159929 DOI: 10.18632/oncotarget.14920].

10. Qiu H, Wang X, Guo R, Liu Q, Wang Y, Yuan Z, Li J, Shi H. HOTAIR rs920778 polymorphism is associated with ovarian cancer susceptibility and poor prognosis in a Chinese population. *Future Oncol* 2017; 13: 347-355 [PMID: 27006321 DOI: 10.2217/ fon-2016-0290].

11. Gong WJ, Yin JY, Li XP, Fang C, Xiao D, Zhang W, Zhou HH, Li X, Liu ZQ. Association of well-characterized lung cancer IncRNA polymorphisms with lung cancer susceptibility and platinum-based chemotherapy response. *Tumour Biol* 2016; 37: 8349-8358 [PMID: 26729200 DOI: 10.1007/s13277-015-4497-5].

12. Wang JZ, Xiang J, Wu LG, Bai YS, Chen ZW, Yin QX, Wang Q, Guo WH, Peng Y, Guo H, Xu P. A genetic variant in long non-coding RNA MALAT1 associated with survival outcome among patients with advanced lung adenocarcinoma: a survival cohort analysis. *BMJ Cancer* 2017; 17: 167 [PMID: 28253859 DOI: 10.1186/s12858-017-3151-6].

13. Kasagi Y, Oki E, Ando K, Ito S, Iuchi T, Sugiyama N, Nakashima Y, Ohgaki K, Saeke H, Mimori K, Maehara Y. The Expression of CCAT2, a Novel Long Noncoding RNA Transcript, and rs6983267 Single-Nucleotide Polymorphism Genotypes in Colorectal Cancers. *Oncology* 2017; 92: 48-54 [PMID: 27875818 DOI: 10.1159/000452143].

14. Hu P, Qiao O, Wang J, Li J, Jin H, Li Z, Jin Y. rs1859168 A gt g C polymorphism regulates HOTTIP expression and reduces risk of pancreatic cancer in a Chinese population. *World J Surg Oncol* 2017; 15: 155 [PMID: 28818070 DOI: 10.1186/s12957-017-1218-0].

15. Xu Q, Chen MY, He CY, Sun LP, Yuan Y. Promoter polymorphisms in trefoil factor 2 and trefoil factor 3 genes and susceptibility to gastric cancer and atrophic gastritis among Chinese population. *Gene* 2013; 529: 104-112 [PMID: 23933418 DOI: 10.1016/j.gene.2013.07.070].

16. Gong Y, He C, Duan Z, Sun L, Xu Q, Xing C, Yuan Y. Association of two ERCC4 tagSNPs with susceptibility to atrophic gastritis and gastric cancer in Chinese. *Gene* 2013; 519: 335-342 [PMID: 23415627 DOI: 10.1016/j.gene.2013.01.059].

17. Vines P, Manuanguarda M, Kavoura FK, Guerrera S, Allione A, Rosa F, Di Gregorio A, Polidoro S, Saletta F, Ioannidis JP, Matullo G. A field synopsis on low-penetration variants in DNA repair genes and cancer susceptibility. *J Natl Cancer Inst* 2009; 101: 24-36 [PMID: 19163638 DOI: 10.1093/jnic/dnp437].

18. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002; 3: 391-397 [PMID: 11988764 DOI: 10.1038/nrg976].

19. Yuan HY, Chiov J, Tseng WH, Liu CH, Liu CK, Lin YJ, Wang HH, Yao A, Chen YT, Hsu CN. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res* 2006; 34: W635-W641 [PMID: 16845089 DOI: 10.1093/nar/gkl236].

20. Xu Q, Yuan Y, Sun LP, Gong YH, Xu Y, Xu XW, Dong NN, Lin GD, Smith PN, Li RW. Risk of gastric cancer is associated with the MUC1 568 A/G polymorphism. *Int J Oncol* 2009; 35: 1313-1320 [PMID: 19885554].

21. Westra HJ, Peters MJ, Esko T, Vaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, Zehnacker A, Zehnacker DV, Veldink JH, Van den Berg LH, Karlainen J, Withoff S, Uitterlinden AG, Hofman A, Rivadeneira F, Hoen PAC, Reinmaa E, Fischer K, Nelis M, Milan Li, Melzer D, Ferrucci L, Singleton AB, Hernandez DG, Nalls MA, Homuth G, Nauck M, Radke D, Völker U, Perola M, Salomaa V, Brody BJ, Suchy-Dickey A, Gharib SA, Enquobahrie DA, Lumley T, Montgomery GW, Makino S, Sutokis H, Herder C, Roden M, Grallert H, Meitinger T, Strauch K, Li Y, Jansen RC, Visscher PM, Grallert H, Meitinger T, Strauch K, Li Y, Jansen RC, Visscher PM, Knight JC, Psaty BM, Ripatti S, Teumer A, Frayling TM, Metspalu A, van Meurs JB, Franke L. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; 45: 1238-1243 [PMID: 24013639 DOI: 10.1038/ng.2756].

22. Li Z, Zhang Z, He Z, Tang W, Li T, Zeng Z, He L, Shi Y. A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHeSis (http://analysis.bio-x.cn). *Cell Res* 2009; 19: 519-523 [PMID: 19290020 DOI: 10.1038/cr.2009.33].

23. Deng X, Zhao Y, Wu X, Song G. Upregulation of CAT2 promotes cell proliferation by repressing the P15 in breast cancer. *Biomed Pharmacother* 2017; 91: 1160-1166 [PMID: 28531944 DOI: 10.1016/j.biopha.2017.05.030].

24. Wu ZJ, Li Y, Wu YZ, Wang Y, Nian WQ, Wang LL, Li LC, Luo HL, Wang DL. Long non-coding RNA CAT2 promotes the breast cancer growth and metastasis by regulating TGF-β signaling.
pathway. *Eur Rev Med Pharmacol Sci* 2017; 21: 706-714 [PMID: 28272713]

25 Cheng Y, Jutooru I, Chadalapaka G, Corton JC, Safe S. The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. *Oncotarget* 2015; 6: 10840-10852 [PMID: 25912306 DOI: 10.18632/oncotarget.3450]

26 Zhang S, Wang W, Liu G, Xie S, Li Q, Li Y, Lin Z. Long non-coding RNA HOTTIP promotes hypoxia-induced epithelial-mesenchymal transition of malignant glioma by regulating the miR-101/ZEB1 axis. *Biomed Pharmacother* 2017; 95: 711-720 [PMID: 28886531 DOI: 10.1016/j.biopha.2017.08.133]

27 Wang Y, Zhang Y, Yang T, Zhao W, Wang N, Li P, Zeng X, Zhang W. Long non-coding RNA MALAT1 for promoting metastasis and proliferation by acting as a ceRNA of miR-144-3p in osteosarcoma cells. *Oncotarget* 2017; 8: 59417-59434 [PMID: 28938647 DOI: 10.18632/oncotarget.19727]

28 Zuo Y, Li Y, Zhou Z, Ma M, Fu K. Long non-coding RNA MALAT1 promotes proliferation and invasion via targeting miR-143-5p in human glioblastoma cell lines. *Biomed Pharmacother* 2017; 95: 711-720 [PMID: 28938647 DOI: 10.18632/oncotarget.19727]

29 Shi B, Wang Y, Liu G, Xu S, Li Q, Li Y, Lin Z. Long non-coding RNA HOTTIP promotes hypoxia-induced epithelial-mesenchymal transition of malignant glioma by regulating the miR-101/ZEB1 axis. *Biomed Pharmacother* 2017; 95: 711-720 [PMID: 28886531 DOI: 10.1016/j.biopha.2017.08.133]

30 Yu Y, Pan S, Liu L, Zhai X, Liu J, Wen J, Zhang Y, Chen J, Shen H, Hu Z. A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. *PLoS One* 2012; 7: e35145 [PMID: 22493738 DOI: 10.1371/journal.pone.0035145]

31 Xu Y, Wang B, Zhang F, Wang A, Du X, Hu P, Zhu Y, Fang Z. Long non-coding RNA CCAT2 is associated with poor prognosis in hepatocellular carcinoma and promotes tumor metastasis by regulating Snail2-mediated epithelial-mesenchymal transition. *Oncol Targets Ther* 2017; 10: 1191-1198 [PMID: 28280353 DOI: 10.2147/OTT.S127100]

32 Zhou N, Si Z, Li T, Chen G, Zhang Z, Qi H. Long non-coding RNA CCAT2 functions as an oncogene in hepatocellular carcinoma, regulating cellular proliferation, migration and apoptosis. *Onco Lett* 2016; 12: 132-138 [PMID: 27347113 DOI: 10.3892/ol.2016.4580]

33 Chen F, Bai G, Li Y, Feng T, Wang L. A positive feedback loop of long noncoding RNA CCAT2 and FOXM1 promotes hepatocellular carcinoma growth. *Am J Cancer Res* 2017; 7: 1423-1434 [PMID: 28744394]

34 Chen L, Yao H, Wang K, Liu X. Long Non-Coding RNA MALAT1 Regulates ZEB1 Expression by Sponging mir-143-3p and Promotes Hepatocellular Carcinoma Progression. *J Cell Biochem* 2017; 118: 4836-4843 [PMID: 28543721 DOI: 10.1002/jcb.26158]

35 Gong WJ, Peng JB, Yin YJ, Li XP, Zheng W, Xiao L, Tan LM, Xiao D, Chen YX, Li X, Zhou HH, Liu QZ. Association between well-characterized lung cancer lncRNA polymorphisms and platinum-based chemotherapy toxicity in Chinese patients with lung cancer. *Acta Pharmacol Sin* 2017; 38: 581-590 [PMID: 28260796 DOI: 10.1038/aps.2016.164]

36 Liu Y, Pan S, Liu L, Zhai X, Liu J, Wen J, Zhang Y, Chen J, Shen H, Hu Z. A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. *PLoS One* 2012; 7: e35145 [PMID: 22493738 DOI: 10.1371/journal.pone.0035145]

37 He C, Vogel U, Ma Y, Qi R, Wang H. HapMap-based study of the DNA repair gene ERCC2 and lung cancer susceptibility in a Chinese population. *Carcinogenesis* 2009; 30: 1181-1185 [PMID: 19406934 DOI: 10.1093/carcin/bgp107]

38 Carlborg O, Haley CS. Epistasis: too often neglected in complex trait studies? *Nat Rev Genet* 2004; 5: 618-625 [PMID: 15266344 DOI: 10.1038/nrg1407]

39 Baker SG, Kramer BS, Sargent DJ, Bonetti M. Biomarkers, subgroup evaluation, and clinical trial design. *Discov Med* 2012; 13: 187-192 [PMID: 22463794]

40 Gupta RM, Musunuru K. Mapping Novel Pathways in Cardiovascular Disease Using eQTL Data: The Past, Present, and Future of Gene Expression Analysis. *Front Genet* 2013; 3: 232 [PMID: 23755065 DOI: 10.3389/fgene.2012.00232]
