

OPEN

# Genetic Polymorphisms in Long Noncoding RNA H19 Are Associated With Susceptibility to Breast Cancer in Chinese Population

Zongjiang Xia, PhD, Rui Yan, MD, Fujiao Duan, MD, Chunhua Song, PhD, Peng Wang, PhD, and Kaijuan Wang, PhD

**Abstract:** H19, a maternally expressed imprinted gene transcribing a long noncoding RNA, has previously been reported to be involved in tumorigenesis and cancer progression. However, the association between the H19 polymorphisms and breast cancer (BC) susceptibility has remained elusive. The aim of this study was to evaluate the associations between 2 H19 haplotype tagging SNPs (rs3741219 T>C, rs217727 C>T) and the risk of breast cancer.

Our study comprised 464 BC patients and 467 cancer-free controls in China. rs3741219 and rs217727 were genotyped with polymerase chain reaction-restriction fragment-length polymorphism (PCR-RFLP) and created restriction site PCR (CRS-RFLP) assays, respectively. False-positive report probability (FPRP) was calculated to test the false-positive association.

On performing univariate analysis, no significant association between H19 polymorphisms (rs3741219 and rs217727) and BC was observed. However, in further stratified analyses, CT+TT genotypes of rs217727 had a significantly lower risk of breast cancer among women with number of pregnancy >2 (OR = 0.79; 95% CI = 0.55–0.97). CT genotype of rs217727 was associated with ER positivity (OR = 2.19; 95% CI = 1.07–4.45) and HER-2 positivity (OR = 1.34; 95% CI = 1.05–2.12). It was proved that our results were less likely to be false positives according to false-positive report probability calculation.

Our findings extend available data on the association of H19 polymorphisms and BC susceptibility. Further validation in large population or cohort studies is needed.

(*Medicine* 95(7):e2771)

**Abbreviations:** 95% CI = 95% confidence interval, BC = breast cancer, CRS-RFLP = created restriction site-restriction fragment-

length polymorphism, ER = estrogen receptor, FPRP = false-positive report probability, HER-2 = human epidermal growth factor receptor-2, IGF2 = insulin-like growth factor 2, lncRNA = long noncoding RNA, OR = odds ratio, PCR-RFLP = polymerase chain reaction-restriction fragment-length polymorphism, PR = progesterone receptor, SNPs = single nucleotide polymorphisms.

## INTRODUCTION

Breast cancer is the most frequently diagnosed malignant tumor and the first leading cause of cancer death among women.<sup>1–3</sup> The development of BC is a complex and multifactorial process, and numerous epidemiological studies have identified several risk factors for BC, involving multiple reproductive factors and susceptibility genes.<sup>4,5</sup> Previous genomics researches have provided evidence to support that single nucleotide polymorphisms (SNPs) in several genes were associated with the risk of BC<sup>6,7</sup> and may contribute to the carcinogenesis of BC.

Long noncoding RNAs (lncRNAs) are defined as transcribed RNA molecules that are >200 nucleotides, lacking open reading frame, and having no obvious protein-coding capacity.<sup>8,9</sup> Several lncRNAs have been reported to be associated with cancers,<sup>10,11</sup> including BC. The handful of aberrantly expressed lncRNAs in different cancers suggests that lncRNAs contribute to carcinogenesis, and deliver functions in controlling cell cycle progression, apoptosis, invasion, and migration.<sup>12</sup> Single nucleotide polymorphisms (SNPs) have been confirmed to have profound effects on gene expression and function, and participate in carcinogenesis. Recently, studies on the effects of SNPs have extended to functional lncRNAs. For example, HOTAIR has been widely identified to participate in tumor pathogenesis, acting as a promoter in colorectal cancer carcinogenesis.<sup>13</sup> HOTAIR rs7958904 CC decreased the risk of colorectal cancer compared with GG genotype<sup>14</sup> and HOTAIR rs920778 TT carriers had increased gastric cancer risk.<sup>15</sup> Increased breast cancer risk was associated with the lincRNA-ENST00000515084 rs12325489 CC and CT genotype when compared with the rs12325489 TT genotype.<sup>16</sup> rs11752942 AG+GG in the lincRNA-uc003opf.1 exon had a significantly reduced risk of esophageal squamous cell carcinoma.<sup>17</sup>

Another lncRNA that may play an important role in carcinogenesis is H19. LncRNA H19 is an imprinted gene, located on human chromosome 11p15.5, encodes a 2.3-kb long, capped, spliced, and polyadenylated noncoding RNA.<sup>18</sup> H19 is abnormally expressed in several tumors, and it acts as either a tumor suppressor,<sup>19,20</sup> or an oncogene.<sup>21</sup> Increasing evidence suggests that H19 genetic variants play important roles in the development of cancers. For example, CT+TT rs2839698 in H19 was associated with significantly increased

Editor: Hao Guo.

Received: October 19, 2015; revised: January 5, 2016; accepted: January 15, 2016.

From the Department of Surgery Medicine, Division of Thoracic Surgery, The First Affiliated Hospital, Zhengzhou University (ZX); Department of Epidemiology and Statistics, College of Public Health, Zhengzhou University (RY, CS, PW, KW); Henan Key Laboratory of Tumor Epidemiology (RY, CS, PW, KW); and Department of Hospital Infection Management, Affiliated Cancer Hospital of Zhengzhou University (FD), Zhengzhou, China.

Correspondence: Chunhua Song, Department of Epidemiology and Statistics, College of Public Health, Zhengzhou University, Zhengzhou, Henan Province, China (e-mail: sch16@zzu.edu.cn).

This study was supported by National Natural Science Foundation of China (81202278) and Medical science and Technology Key Projects of Henan province (201303005).

ZX and RY contributed equally to this study.

The authors have no conflicts of interest to disclose.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ISSN: 0025-7974

DOI: 10.1097/MD.0000000000002771

gastric cancer risk.<sup>22</sup> A significantly decreased risk of bladder cancer was found for H19 rs2839698 TC carriers.<sup>23</sup> However, according to our latest knowledge, no research has been executed to evaluate the H19 polymorphism and the risk of BC. On the basis of the above description, we hypothesized that functional SNPs in H19 might have association with the BC risk. We genotyped 2 H19 haplotype tagging SNPs in a population-based case-control study comprising 464 BC patients and 467 age frequency matched controls from China. The associations between the H19 SNPs and breast cancer risk were investigated by molecular epidemiology.

## MATERIALS AND METHODS

### Study Population

There were 464 BC patients and 467 cancer-free controls in our study. All subjects were genetically unrelated ethnic Han Chinese women. Patients with newly pathologically confirmed breast cancer were recruited from the First Affiliated Hospital of Zhengzhou University and the Third Affiliated Hospital of Zhengzhou University between 2014 and 2015. We then selected 467 healthy participants, who are genetically unrelated with cases, free of any cancer and having no the history of chronic diseases, as our control group, from a database consisting of >20,000 subjects participated community-based chronic diseases program in Henan province during the same time period that the cases were collected. All healthy controls were frequency matched to the cancer patients on the basis of their age ( $\pm 2$  years).

A structured questionnaire was used to retrieve reproductive information on the subjects, such as age, age of menarche, premenopausal or postmenopausal, age of menopause, number of pregnancy, number of abortion, breast-feeding history for born baby (yes, no), and family history of BC in first-degree relatives (yes, no). The questionnaires were administered by trained interviewers through face-to-face interviews, and informed consent was obtained from each study participant. Pathological data of BC patients, including of the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER-2) status, were collected by immunohistochemistry (IHC) from the patients' pathology reports. The study was approved by the ethical review committee of Zhengzhou University Committee for Medical and Health Research Ethics.

### DNA Extraction

Venous blood (5 mL) was collected into a test tube containing ethylene diamine tetra acetic acid (EDTA) from each BC patient and healthy women control. Genomic DNA was

extracted from peripheral blood samples using the DNA Extraction Kit of TIANGEN BIOTECH (Beijing) according to the manufacturer's instructions. The extracted DNA was stored at  $-80^{\circ}\text{C}$  until use.

### SNPs Selection and Genotyping

Haplotype-tagging SNPs in lncRNA H19 in the chromosomal region 11p15.5 were selected by using Haploview software 4.2, based on the Chinese Han population data of HapMap (HapMap Data Rel 28 Phase II + III, August 10, on NCBI B36 assembly, dbSNP b126). We searched tagging SNPs that captured all the known common SNPs located in the chromosome locus transcribed into H19 and its flanking region (2000 bp upstream and 2000 bp downstream, respectively). Finally, 2 tagging SNPs (rs3741219 T>C, rs217727 C>T) that located in the H19 were selected, with a pairwise correlation  $r^2 > 0.8$ . The genotyping of rs3741219 was determined by polymerase chain reaction-restriction fragment-length polymorphism (PCR-RFLP), whereas rs217727 was genotyped with created restriction site PCR (CRS-RFLP) assays.

The primers designed by Primer 6.0 software used for PCR amplification were 5'-CCCCCTGCGGCGGACGGTTGA-3' (forward) and 5'-GGCGTAATGGAATGCTTGAA-3' (reverse) for rs3741219, and 5'-ACTCAGGAATCGGCTCTG-GAAGGTG-3' (forward) and 5'-GATGTGGTGGCTGGTGGT-CAACGGT-3' (reverse) for rs217727 (Table 1). PCR primers were further verified by NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi/>) to assess the possibility of amplification of any nonspecific DNA sequences and synthesized commercially.

For each sample, PCR amplification was performed in a final volume of 30  $\mu\text{L}$ , which contained 15  $\mu\text{L}$  2  $\times$  Tap PCR MasterMix, 0.5  $\mu\text{L}$  each primer (10  $\mu\text{M}$ ), 50 ng DNA, and 13  $\mu\text{L}$  deionized water. Thermocycling conditions of PCR were as follows: initial denaturation at 95  $^{\circ}\text{C}$  for 5 min, 35 cycles of PCR consisting of denaturation at 94  $^{\circ}\text{C}$  for 30 s, optimal annealing temperature (61.0  $^{\circ}\text{C}$  for rs3741219, and 63.0  $^{\circ}\text{C}$  for rs217727) for 45 s and extension at 72  $^{\circ}\text{C}$  for 45 s, and final extension step of 72  $^{\circ}\text{C}$  for 5 min. In addition, the restriction enzyme HhaI and RsrII (Fermentas, Canada), selected by WATCUT website (<http://watcut.uwaterloo.ca/watcut/watcut/template.php>), were used for genotyping of rs3741219 and rs217727, respectively.

All analyses were performed without knowledge of the case or control status for quality control. Totally 10% of the study populations were randomly selected from both cases and controls to confirm the genotyping results by the direct sequencing (BGI Sequencing, Beijing). The results of confirmation were found to be 100% concordant.

TABLE 1. PCR Information of the 2 SNPs

SNP	Chr:position	MAF*	Genotyping Assay	Annealing Tm ( $^{\circ}\text{C}$ )	Primers
rs3741219 T>C	Chr11:995389	C = 0.292	PCR-RFLP	61.0	Sense: CCCCCTGCGGCGGACGGTTGA Antisense: GGCGTAATGGAATGCTTGAA
rs217727 C>T	Chr11:1995678	T = 0.320	CRS-RFLP	63.0	Sense: ACTCAGGAATCGGCTCTGGAAGGTG Antisense: GATGTGGTGGCTGGTGGTCAACGGT

MAF = minor allele frequency.

\*Based on the Chinese Han population data of the international HapMap project.

**TABLE 2.** Characteristics of Breast Cancer Cases and Cancer-Free Controls

Characteristics	Cases (%) (n = 464)	Controls (%) (n = 467)	P*	OR (95% CI)
Age (mean ± SD)	48.44 ± 10.29	48.91 ± 9.99	0.60	
Age at menarche (mean ± SD)	14.45 ± 1.79	14.25 ± 1.75	0.49	
Age at menopause				
≤50	118 (71.95)	149 (81.42)		1
>50	43 (26.22)	30 (16.39)	0.03	1.81 (1.07,3.06)
Unknown	3 (1.83)	4 (2.19)		
Menstrual history				
Premenopausal	249 (53.66)	259 (55.46)		1
Postmenopausal	164 (35.34)	183 (39.19)	0.63	0.93 (0.71,1.23)
Unknown	51 (10.90)	25 (5.35)		
No. of pregnancy				
0–2	176 (37.93)	209 (44.75)		1
≥3	230 (49.57)	216 (52.67)	0.09	1.26 (0.96,1.66)
Unknown	58 (12.50)	42 (8.98)		
No. of abortion				
0–2	357 (76.94)	375 (80.30)		1
≥3	47 (10.13)	43 (9.21)	0.58	1.15 (0.74,1.78)
Unknown	60 (12.93)	49 (10.49)		
Breast-feeding				
No	21 (4.53)	31 (6.64)		1
Yes	404 (87.07)	399 (85.44)	0.20	1.50 (0.84,2.65)
Unknown	39 (8.40)	37 (7.92)		
Family history of BC				
No	407 (87.72)	419 (89.72)		1
Yes	19 (4.09)	12 (2.57)	0.21	1.63 (0.78,3.40)
Unknown	38 (8.19)	36 (7.71)		

BC = breast cancer, CI = confidence interval, OR = odds ratio, SD = standard deviation.

\* Two-sided  $\chi^2$  test,  $P < 0.05$  was considered statistically significant.

### Analysis

Hardy–Weinberg equilibrium (HWE) was evaluated by a goodness-of-fit chi-squared ( $\chi^2$ )-test to compare the observed genotype frequencies with the expected ones among the cancer-free control subjects. The differences in the distributions of age, reproductive variables, as well as the SNPs genotype frequencies between BC cases and controls, were appraised by using Student’s *t* test for continuous variables, whereas the chi-squared ( $\chi^2$ ) test was used for categorical variables. The relationship between genetic variants and case–control status was examined using unconditional logistic regression analysis (Stepwise selection), with adjustments for age, age at menarche, status of menopausal, number of pregnancy, number of abortion, breast-feeding history for born baby, family history of BC in first-degree relatives. Odds ratio (OR) and its corresponding 95% confidence interval (95% CI) were calculated to estimate the strength of association between the H19 polymorphisms and risk of BC. Furthermore, the data were stratified by age and reproductive factors to evaluate the stratum variable-related ORs among various H19 SNPs. Haplotype analysis was conducted using the online SHEsis (<http://analysis.bio-x.cn/myAnalysis.php>). For all significant genetic effects observed in our study, the false-positive associations were calculated by FPRP with prior probabilities of 0.001, 0.01, 0.1, and 0.25. The OR was set at 1.5 under dominant genetic model, and a probability  $< 0.5$  was considered as noteworthy. Statistical analysis was

performed by using SPSS 16.0 software package (SPSS Inc., Chicago, IL). All tests were 2 sided and a *P* value  $< 0.05$  was considered to be statistically significant.

## RESULTS

### Population Characteristics

The detailed demographic and reproductive characteristics of study subjects were illustrated in Table 2. The mean age was  $48.44 \pm 10.29$  and  $48.91 \pm 9.99$  years for BC cases and healthy controls, respectively. As expected, the mean age for 2 groups paired quite well. No significant differences between cases and control groups with respect to most of the baseline characteristic factors, including age, age at menarche and menopause, menstrual history, No. of abortion, breast-feeding, and family history. Only women with older age of menopause ( $>50$ ) was found to be associated with significant increased BC risk (OR = 1.81; 95% CI = 1.07–3.06).

### H19 Genotypes and BC Risk

The genotype distribution of 2 tagSNPs and their associations with BC risk were shown in Table 3. No significant deviations from Hardy–Weinberg equilibrium for the 2 polymorphisms was found among the 467 healthy controls ( $P > 0.05$ ). Genotyping results showed that only rs217727 CT was significantly associated with decreased BC risk

**TABLE 3.** Genotype Among Cases and Controls and Their Association With BC Risk

Genotype	Cases (%) (n = 464)	Controls (%) (n = 467)	P*	P†	OR (95% CI)‡	P‡	OR (95% CI)‡
rs374129							
TT	238 (51.29)	245 (52.46)			1		1
TC	186 (40.09)	182 (38.97)		0.71	1.05 (0.80, 1.38)	0.31	1.23 (0.83, 1.83)
CC	40 (8.62)	40 (8.57)	0.458	0.90	1.03 (0.64, 1.65)	0.54	1.25 (0.61, 2.54)
TC+CC	226 (48.71)	222 (47.54)		0.72	1.05 (0.81, 1.36)	0.21	1.23 (0.85, 1.79)
T	662 (71.34)	672 (71.95)			1		
C	266 (28.66)	262 (28.05)		0.29	1.10 (0.92, 1.31)		
rs217727							
CC	160 (34.48)	139 (29.76)			1		1
CT	156 (33.62)	212 (45.40)		0.01	0.64 (0.47, 0.87)	0.57	0.88 (0.56, 1.37)
TT	148 (31.90)	116 (24.84)	0.052	0.55	1.11 (0.80, 1.55)	0.17	1.40 (0.87, 2.27)
CT+TT	304 (65.52)	328 (70.24)		0.14	0.81 (0.61, 1.06)	0.75	1.07 (0.72, 1.59)
C	476 (51.29)	490 (52.46)			1		
T	452 (48.71)	444 (47.54)		0.61	1.05 (0.88, 1.2)		

BC = breast cancer, CI = confidence interval, OR = odds ratio.

\* P value of Hardy–Weinberg equilibrium in controls.

† Chi-square test for genotype distributions between cases and controls.

‡ Data were calculated by logistic regression analysis (stepwise selection) with adjusted for age, age at menarche, menopausal status, number of pregnancy and abortion, breast-feeding status, family history of BC in first-degree relatives.

(OR = 0.64; 95% CI = 0.47–0.87). However, the significant decreased BC risk was lost in multivariate analysis (OR = 0.88; 95% CI = 0.56–1.37), after adjusted for age, age at menopause, menopausal status, number of pregnancy and abortion, breast-feeding status and family history of BC in first-degree relatives. Neither the variant TC, CC, TC+CC genotypes nor C allele of rs3741219 had significant association with risk of BC.

**Haplotype Analyses**

We further evaluated the combined effect of the 2 polymorphisms on the risk of BC by using the haplotype analysis. A total of 4 haplotypes were derived from the observed genotypes (Table 4), of which T<sub>rs3741219</sub>T<sub>rs217727</sub> was the most common haplotype in cases and controls. No significant association with BC risk was observed for these 4 haplotypes.

**Combined Effect of 2 SNPs**

We then categorized the 2 SNPs into the number of combined variant alleles to further analyze their possible joint effect

**TABLE 4.** Haplotype Analysis of rs374129 and rs217727 Polymorphism Sites in H19

Haplotype*	Cases (%)	Controls (%)	χ <sup>2</sup>	P	OR (95% CI)
C C	159.75 (17.2)	166.43 (17.8)	0.117	0.73	0.96 (0.76, 1.22)
C T	106.25 (11.4)	95.57 (10.2)	0.713	0.40	1.13 (0.85, 1.52)
T C	316.25 (34.1)	323.57 (34.6)	0.066	0.79	0.98 (0.81, 1.18)
T T	345.75 (37.3)	348.43 (37.3)	0.000	0.98	1.00 (0.83, 1.20)

BC = breast cancer, CI = confidence interval, OR = odds ratio.

\* SNPs sequence: rs374129 and rs217727.

and potential locus–locus interaction on BC risk. As shown in Table 5, individuals with 1 mutation allele had 39% decreased BC risk when compared to individuals with 0 mutation allele (OR = 0.61; 95% CI = 0.40–0.91). No statistical risk for BC in other subgroups and no increased dose-dependent manner were observed on the combined effect of the 2 SNPs.

**Stratified Analysis of SNP Genotypes and BC Risk**

Next, we performed a stratified analysis using logistic regression to evaluate the associations between H19 polymorphisms and BC risk. As indicated in Table 6, CT+TT genotypes of rs217727 had a significantly lower risk of breast cancer among women with no. of pregnancy >2 (OR = 0.79; 95% CI = 0.55–0.97).

**TABLE 5.** Combined Effect of the 2 SNPs on Breast Cancer

Combined SNPs*	Cases (%)	Controls (%)	χ <sup>2</sup>	P	OR (95% CI)
0	78 (16.81)	61 (13.06)			1
1	133 (28.66)	172 (36.83)	5.99	0.01	0.61 (0.40, 0.91)
2	181 (39.01)	175 (37.47)	1.11	0.29	0.81 (0.55, 1.20)
3	65 (14.01)	52 (11.14)	0.01	0.93	0.98 (0.60, 1.60)
4	7 (1.51)	7 (1.50)	0.19	0.66	0.78 (0.26, 2.35)
			0.33†	0.56†	
Total	464	467			

CI = confidence interval, OR = odds ratio, SNPs = single nucleotide polymorphisms.

\* The SNPs: rs374129, rs217727. “0” represents rs3741219 TT with rs217727 CC. “1” represents rs3741219 TC with rs217727 CC, and rs3741219 TT with rs217727 CT. “2” represents rs3741219 TT with rs217727 TT, rs3741219 CC with rs217727 CC, and rs3741219 TC with rs217727 CT. “3” represents rs3741219 CC with rs217727 CT, and rs3741219 TC with rs217727 TT. “4” represents rs3741219 CC with rs217727 TT.

† Trend chi-square and P values.

**TABLE 6.** Stratification Analysis of the 2 SNPs Polymorphisms and BC Susceptibility

Variables	rs3741219 (TT/TC+CC)			P <sup>†</sup>	rs217727 (CC/CT+TT)			P <sup>†</sup>
	Case	Control	OR (95% CI)*		Case	Control	OR (95% CI)*	
Age								
≤50	153/138	142/140	1.16 (0.71,1.492)	0.55	99/192	88/194	1.03 (0.61,1.73)	0.91
>50	85/88	103/82	1.41 (0.81,2.47)	0.23	61/112	51/134	1.03 (0.56,1.88)	0.92
Age at menarche								
≤14	130/120	150/135	1.26 (0.78,2.03)	0.33	81/169	80/205	0.85 (0.52,1.41)	0.54
>14	81/81	82/74	1.25 (0.68,2.30)	0.47	59/103	53/103	1.52 (0.79,2.95)	0.21
Age at menopause								
≤50	59/59	87/65	1.53 (0.78,3.00)	0.22	38/80	39/113	1.19 (0.56,2.50)	0.65
>50	24/17	23/111	1.37 (0.49,3.87)	0.55	19/28	7/25	0.63 (0.21,1.87)	0.41
Menopause								
Premenopausal	130/119	128/131	1.13 (0.68,1.88)	0.64	83/166	87/172	1.12 (0.66,1.90)	0.68
Postmenopausal	82/82	105/78	1.33 (0.75,2.34)	0.33	57/107	46/137	1.06 (0.58,1.95)	0.86
No. of pregnancy								
≤2	87/89	102/107	1.04 (0.56,1.95)	0.90	57/119	63/146	0.88 (0.45,1.71)	0.72
>2	120/110	122/94	1.28 (0.80,2.06)	0.31	81/149	66/150	0.79 (0.55,0.97)	0.03
No. of abortion								
≤2	184/173	204/171	1.33 (0.89,1.98)	0.17	123/234	118/257	1.00 (0.66,1.51)	0.98
>2	22/25	19/24	1.01 (0.35,2.95)	0.98	15/34	9/34	1.68 (0.47,6.01)	0.42
Breast-feeding								
No	11/10	19/12	1.86 (0.71,4.87)	0.21	4/17	10/21	1.41 (0.55,3.63)	0.47
Yes	206/198	209/190	1.25 (0.85,1.84)	0.27	139/265	118/281	0.99 (0.66,1.50)	0.98
Family history								
No	211/196	215/204	1.24 (0.84, 1.82)	0.28	141/266	129/290	1.06 (0.71,1.60)	0.77
Yes	9/10	8/4	0.67 (0.19,4.93)	0.69	7/12	3/9	2.33 (0.20,7.56)	0.50

CI = confidence interval, OR = odds ratio.

\*Adjusted for age, age at menarche, menopausal status, number of pregnancy and abortion, breast-feeding status, family history of BC in first-degree relatives (the stratified factor in each stratum excluded).

†P value for the heterogeneity test.

### Receptor Status and BC risk

We further demonstrated the association of rs3741219 and rs217727 polymorphism genotypes with the clinicopathological features in Table 7, including ER status, PR status and HER-2 status. In the case only analysis, compared with CC genotype of rs217727, CT genotype was associated with ER positivity (OR = 2.19; 95 % CI = 1.07–4.45) and HER-2 positivity (OR = 1.34; 95 % CI = 1.05–2.12). No significant association was found between rs3741219 variants and ER, PR, HER-2 status.

### FPRP Calculation

We further calculated the FPRP values for all significant genetic effects observed in our study (Table 8). When we set the prior probability at 0.1, the association between rs217727 CT+TT genotypes and breast cancer for subjects with no. of pregnancy >2 was still significant (FPRP = 0.188). The positive association between individuals with 1 mutation allele and breast cancer was still noteworthy (FPRP = 0.295).

### DISCUSSION

In the present molecular epidemiological study, we investigated the effects of 2 potentially functional polymorphisms (rs217727 and rs3741219) presented in H19 on BC

susceptibility through a population-based case-control study in a Chinese Han population. To the best of our knowledge, this is the first study to examine the role of H19 polymorphism in BC tumorigenesis. No significant association between H19 polymorphisms (rs3741219 and rs217727) and BC risk was observed in our study. However, in further stratified analyses, CT+TT genotypes of rs217727 had a significantly lower risk of breast cancer among women with no. of pregnancy >2 (OR = 0.79; 95% CI = 0.55–0.97). CT genotype of rs217727 was significant associated with ER positivity (OR = 2.19; 95 % CI = 1.07–4.45) and HER-2 positivity (OR = 1.34; 95 % CI = 1.05–2.12). Our findings support the hypothesis that potentially functional SNPs of H19 may play a role in the etiology of BC and H19 genetic mutation may have interaction with ER and HER-2 during human breast tumorigenesis.

Previous studies have indicated that lncRNAs could deliver functions in controlling cell cycle progression, apoptosis, invasion, and migration. A genetic variation on lncRNA may change the lncRNA structure, affect the stability of the lncRNA, influence the level of lncRNA expression, and contribute to carcinogenesis. For example, a study by Wu et al, reported that when compared with the A allele, the rs11752942G allele could markedly attenuate the level of lincRNA-uc003opf.1 by binding micro-RNA-149, and affect esophageal squamous cell proliferation and tumor growth.<sup>17</sup> Li et al founded that rs12325489 C to T base change could disrupt the binding site for miRNA-370,

TABLE 7. The Associations Between 2 SNPs and ER, PR and HER-2 Status of Breast Cancer Patients

Genotype	ER			PR			HER-2		
	Negative (n = 138) %	Positive (n = 277) %	OR (95% CI)* P*	Negative (n = 173) %	Positive (n = 241) %	OR (95% CI)* P*	Negative (n = 115) %	Positive (n = 287) %	OR (95% CI)* P*
rs374129									
TT	78 (56.52)	136 (49.10)	1	90 (52.02)	124 (51.45)	1	56 (48.70)	154 (53.66)	1
TC	49 (35.51)	115 (41.52)	0.72	71 (41.04)	92 (38.17)	0.84 (0.46,1.51)	52 (45.22)	106 (36.93)	0.79
CC	11 (7.97)	26 (9.38)	0.41	12 (6.94)	25 (10.37)	1.32 (0.45,3.84)	7 (6.08)	27 (9.41)	0.41
TC+CC	60 (43.78)	141 (50.90)	0.55	83 (29.96)	117 (48.54)	0.74 (0.52,1.09)	59 (51.30)	133 (46.34)	0.97
rs217727									
CC	54 (39.13)	88 (31.77)	1	61 (22.02)	81 (33.61)	1	43 (37.39)	93 (32.40)	1
CT	37 (26.81)	102 (36.82)	0.03	55 (19.86)	83 (34.44)	1.60 (0.78,3.30)	33 (28.70)	100 (34.84)	0.03
TT	47 (34.06)	87 (31.41)	0.84	57 (20.58)	77 (31.95)	0.90 (0.46,1.75)	39 (33.91)	94 (32.75)	0.78
CT+TT	84 (60.87)	189 (68.23)	0.17	112 (40.43)	160 (66.39)	1.21 (0.67,2.18)	72 (62.60)	194 (67.60)	0.34

BC = breast cancer, CI = confidence interval, ER = estrogen receptor, HER-2 = human epidermal growth factor receptor, OR = odds ratio, PR = progesterone receptor.  
\* Adjusted for age, age at menarche, menopausal status, number of pregnancy and abortion, breast-feeding status, family history of BC in first-degree relative.

influencing lincRNA-ENST00000515084 transcriptional activity and affecting breast cancer cell proliferation.<sup>16</sup> The H19 gene belongs to a highly conserved imprinted gene cluster that play important roles in embryonal development and growth control. The cluster also contains the nearby reciprocally imprinted gene for insulin-like growth factor 2 (IGF2).<sup>24,25</sup> H19/IGF2 are closely linked and coordinately regulated by an intergenic control region (ICR) and a common enhancer region. In human, binding of the CTCF insulator protein to the nonmethylated ICR promotes H19 expression from the maternal allele, whereas IGF2 is expressed from the methylated allele. Recently, several H19 polymorphisms have been studied to be associated with diseases. For example, Gao et al found that rs217727 C to T variant was associated with increased coronary artery disease risk (CAD) (CT + TT vs CC: OR = 1.49; 95% CI = 1.21–1.83), whereas the rs2067051 G to A variant was associated with reduced CAD risk (GA + AA vs GG: OR = 0.77; 95% CI = 0.62–0.95).<sup>26</sup> Another research shown that maternal H19 rs217727 TT genotype was associated with a higher birth weight.<sup>27</sup> A population-based case-control study comprising 177 bladder cancer patients and 204 controls reported that rs2839698 TC genotype was statistical significantly associated with decreased risk of bladder cancer compared with TT carriers (OR = 0.60; 95% CI = 0.36–0.99). And a borderline significantly decreased risk of bladder cancer was found for rs217727 T allele carriers compared with C allele carriers (OR = 0.74; 95% CI = 0.51–1.06).<sup>28</sup> However, a previous study by Yang et al evaluated the effects of 4 independent SNPs (H19 rs217727, rs2839698, rs3741216, and rs3741219) on gastric cancer risk in the Chinese Han population and suggested that the T allele of rs217727 was associated with higher risk of gastric cancer in both codominant (OR = 1.34; 95% CI = 1.03–1.73) and recessive (OR = 1.48; 95% CI = 1.04–2.10) inheritance genetic models.<sup>22</sup> Individuals with rs2839698 TT genotype had a significant increased risk of gastric cancer (OR = 1.64; 95% CI = 1.05–2.57) and carriers with rs2839698 T allele was significantly associated with higher H19 mRNA expression in cancer-free controls,<sup>22</sup> suggesting a potential functional impact of H19 SNP on mRNA levels and the susceptibility to gastric cancer. In this population-based case-control study, we found a negative association between individuals with CT genotype of rs217727 and BC susceptibility (OR = 0.64; 95% CI = 0.47–0.87). However, the significant association was lost in the logistic regression analysis (OR = 0.88; 95% CI = 0.56–1.37). Our results may require confirmation in larger prospective studies and different ethnicities in future.

By using the software program ([http:// bioinfo.life.hust.edu.cn/lincRNASNP/](http://bioinfo.life.hust.edu.cn/lincRNASNP/)), no miRNAs that combined with the C/T polymorphism rs217727 was found. However, we predicted that the T to C base change at rs3741219 T>C may create hsa-miR-1539 microRNA (miRNA) binding sites on H19. A single nucleotide polymorphism of lincRNA may change the lincRNA structure, affect the stability and influence the interaction of miRNA-lincRNA. Increasing evidence shows that lincRNAs can be directly regulated by miRNAs.<sup>22</sup> It is unclear that whether the conversion of T > C in the rs3741219 polymorphism in the H19 could affect the miRNA-lincRNA interactions in breast cancer cells, and further studies are needed to explore the specific mechanisms.

H19 is an estrogen-regulated long noncoding RNA<sup>29</sup> and harbors the microRNA-675 in its first exon. Previous investigations have indicated that the controlled release of miR-675 from H19 could inhibit estrogen-induced proliferation of ERα+

**TABLE 8.** FPRP Values for Associations Between BC risk, ER, and Genotypes in Stratified Factors

Genotype	Stratified Factors	BC, ER, HER-2	Positive OR (95% CI)*	P <sup>†</sup>	Prior Probability			
					0.25	0.1	0.01	0.001
rs217727	Age at menopause>50	BC	1.81 (1.07, 3.06)	0.03	0.250	0.498	0.916	0.991
	CT /CC	All subjects	BC	0.64 (0.47, 0.87)	0.01	0.032	0.090	0.522
	CT /CC	All cases	ER	2.19 (1.07, 4.45)	0.03	0.380	0.684	0.953
	CT /CC	All cases	HER-2	1.34 (1.05,2.12)	0.03	0.480	0.735	0.968
	CT+TT/CC	No. of pregnancy>2	BC	0.79 (0.55,0.97)	0.03	0.072	0.188	0.718
No. of mutation allele	1/0	All subjects	BC	0.61 (0.40, 0.91)	0.01	0.122	0.295	0.822

BC = breast cancer, CI = confidence interval, ER = estrogen receptor, HER-2 = human epidermal growth factor receptor, OR = odds ratio.

\* The crude OR.

† The logistic regression analysis of the genotype frequency distributions.

breast cancer cells.<sup>30</sup> Compared with ERa-breast tumors, ERa+ breast tumors exhibited a higher expression of H19.<sup>31</sup> In our study, we investigated the association between the H19 polymorphism and estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER-2) for the first time. In the case only analysis, rs217727CT genotype were associated with ER positivity (OR = 2.19; 95 % CI = 1.07–4.45) and HER-2 positivity (OR = 1.34; 95 % CI = 1.05–2.12), which was a novel finding. It is possible to propose the hypothesis that H19 genetic mutation may have interaction with ER and HER-2 during human breast tumorigenesis. It is worthwhile to better understand and clarify the underlying molecular mechanisms in future studies.

Haplotype was accepted as a group of correlated SNPs that are located in the same homologous chromosome and passed on to descendants together as a whole.<sup>32</sup> Haplotype contained multiple SNPs information, and the statistical power of haplotype analysis was stronger than analysis of single SNP.<sup>33</sup> This study analyzed haplotypes of rs3741219 and rs217727, and no significant positive or negative association with BC risk was found in the 4 haplotypes. This may be because the borderline significant increased BC risk of rs3741219 could exhibit its restraint upon the significant reduced BC risk of rs217727. In order to find more credible evidence on the association between H19 haplotype and BC risk, further studies on larger populations are required.

The results of molecular epidemiology studies were always accompanied by high probability of false positive.<sup>34–36</sup> In order to test the false-positive associations, we subsequently calculated the FPRP for all significant genetic effects observed in our study. The association between rs217727 CT+TT genotypes in subjects with no. of pregnancy >2 and breast cancer, and the positive association between individuals with 1 mutation allele and breast cancer were still noteworthy, which imply the functional SNPs in H19 might be involved in the breast cancer development with a high likelihood.

There were some strengths of this study that should be noted. First, although the role of H19 in multiple diseases has been reported, this is the first study investigating genetic variation of H19 in relation to breast cancer susceptibility. Second, the whole controls in our study were selected from people in a large sampling survey based on community, not from hospital, diminishing the effect of selection bias. To avoid the prevalence–incidence bias, the cases in our study were all

newly pathological diagnosed. The controls and the cases were matched on age, and the baseline characteristic distributions were similar between the 2 groups. Therefore, we believed that selection bias was not substantial and not likely to influence the analyses of our study. Furthermore, for all significant genetic effects observed in our study, we further adjusted potential confounding factors in final analyses and calculated the FPRP. It is proved that our results were less likely to be false positives or false negative according to the FPRP results. However, some limitations of the present study should be mentioned. The statistical power of the present study may be restricted by the sample size and large proportion of missing data for some variable. And our study did not include detection of the influence of the H19 variants on its expression and activity. Therefore, further studies are worthwhile to validate whether H19 SNP could affect the expression of H19 and the etiology of BC.

In conclusion, the present population-based case–control study provided the first evidence that H19 rs217727 might play an interaction with ER positive and HER-2 positivity during human breast tumorigenesis. H19 genetic polymorphisms (rs3741219 and rs217727) may have no association with breast cancer risk after adjusted for reproductive variables. The study may require confirmation in larger prospective studies and different ethnicities with a gene expression functional assay.

**REFERENCES**

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *Ca Cancer J Clin.* 2011;61:69–90.
2. DeSantis C, Ma J, Bryan L, et al. Breast cancer statistics, 2013. *CA Cancer J Clin.* 2014;64:52–62.
3. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. *Lancet Oncol.* 2014;15:e279–289.
4. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. *Lancet Oncol.* 2001;2:133–140.
5. Zhang B, Beeghly-Fadiel A, Long J, et al. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncol.* 2011;12:477–488.
6. Bradbury AR, Olopade OI. Genetic susceptibility to breast cancer. *Rev Endocr Metab Disord.* 2007;8:255–267.

7. Sapkota Y, Mackey JR, Lai R, et al. Assessing SNP–SNP Interactions among DNA repair, modification and metabolism related pathway genes in breast cancer susceptibility. *PLoS One*. 2013;8:e64896.
8. Bertone P, Stolc V, Royce TE, et al. Global identification of human transcribed sequences with genome tiling arrays. *Science*. 2004;306:2242–2246.
9. Mattick JS. The functional genomics of noncoding RNA. *Science*. 2005;309:1527–1528.
10. Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 2010;464:1071–1076.
11. Gu Y, Chen T, Li G, et al. LncRNAs: emerging biomarkers in gastric cancer. *Future Oncol*. 2015;11:2427–2441.
12. Zhao J, Liu Y, Zhang W, et al. Long non-coding RNA Linc00152 is involved in cell cycle arrest, apoptosis, epithelial to mesenchymal transition, cell migration and invasion in gastric cancer. *Cell Cycle*. 2015;14:1–12.
13. Sun H, Wang G, Peng Y, et al. H19 lncRNA mediates 17beta-estradiol-induced cell proliferation in MCF-7 breast cancer cells. *Oncol Rep*. 2015;33:3045–3052.
14. Xue Y, Gu DY, Ma GX, et al. Genetic variants in lncRNA HOTAIR are associated with risk of colorectal cancer. *Mutagenesis*. 2015;30:303–310.
15. Pan W, Liu L, Wei J, et al. A functional lncRNA HOTAIR genetic variant contributes to gastric cancer susceptibility. *Mol Carcinog*. 2015;55:90–96.
16. Li N, Zhou P, Zheng J, et al. A polymorphism rs12325489C > T in the lincRNA-ENST00000515084 exon was found to modulate breast cancer risk via GWAS-based association analyses. *Plos One*. 2014;9:e98251.
17. Wu HC, Zheng J, Deng JQ, et al. A genetic polymorphism in lincRNA-uc003opf.1 is associated with susceptibility to esophageal squamous cell carcinoma in Chinese populations. *Carcinogenesis*. 2013;34:2908–2917.
18. Brannan CI, Dees EC, Ingram RS, et al. The product of the H19 gene may function as an RNA. *Mol Cell Biol*. 1990;10:28–36.
19. Yoshimizu T, Miroglio A, Ripoché MA, et al. The H19 locus acts in vivo as a tumor suppressor. *Proc Natl Acad Sci USA*. 2008;105:12417–12422.
20. Zhang L, Yang F, Yuan JH, et al. Epigenetic activation of the MiR-200 family contributes to H19-mediated metastasis suppression in hepatocellular carcinoma. *Carcinogenesis*. 2013;34:577–586.
21. Matouk IJ, DeGroot N, Mezan S, et al. The H19 non-coding RNA is essential for human tumor growth. *Plos One*. 2007;2:e845.
22. Yang C, Tang R, Ma X, et al. Tag SNPs in long non-coding RNA H19 contribute to susceptibility to gastric cancer in the Chinese Han population. *Oncotarget*. 2015;6:15311–15320.
23. Verhaegh GW, Verkleij L, Vermeulen SH, et al. Polymorphisms in the H19 gene and the risk of bladder cancer. *Eur Urol*. 2008;54:1118–1126.
24. Takai D, Gonzales FA, Tsai YC, et al. Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. *Hum Mol Genet*. 2001;10:2619–2626.
25. Wilkin F, Paquette J, Ledru E, et al. H19 sense and antisense transgenes modify insulin-like growth factor-II mRNA levels. *Eur J Biochem/FEBS*. 2000;267:4020–4027.
26. Gao W, Zhu M, Wang H, et al. Association of polymorphisms in long non-coding RNA H19 with coronary artery disease risk in a Chinese population. *Mutat Res*. 2015;772:15–22.
27. Hewage AS, Jayanthiny P, Tennekoon KH, et al. H19 rs217727 genotype and IGF-1/intron-2 dinucleotide CT repeat polymorphism are independently associated with birth weight. *Endocrine*. 2015;48:1010–1012.
28. Verhaegh GW, Verkleij L, Vermeulen SHM, et al. Polymorphisms in the H19 gene and the risk of bladder cancer. *Eur Urol*. 2008;54:1118–1126.
29. Adriaenssens E, Lottin S, Dugimont T, et al. Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene*. 1999;18:4460–4473.
30. Keniry A, Oxley D, Monnier P, et al. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat Cell Biol*. 2012;14:659–665.
31. Basak P, Chatterjee S, Weger S, et al. Estrogen regulates luminal progenitor cell differentiation through H19 gene expression. *Endocr-Relat Cancer*. 2015;22:505–517.
32. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296:2225–2229.
33. Akey J, Jin L, Xiong MM. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J Hum Genet*. 2001;9:291–300.
34. Colhoun HM, McKeigue PM, Smith GD. Problems of reporting genetic associations with complex outcomes. *Lancet*. 2003;361:865–872.
35. Ioannidis JPA, Ntzani EE, Trikalinos TA, et al. Replication validity of genetic association studies. *Nat Genet*. 2001;29:306–309.
36. Sterne JA, Davey Smith G. Sifting the evidence—what’s wrong with significance tests? *BMJ*. 2001;322:226–231.