



Association of long non-coding RNA HOTAIR and MALAT1 variants with cervical cancer risk in Han Chinese women

Meiqun Jia^{1,2Δ}, Lulu Ren^{1Δ}, Lingmin Hu^{1,3}, Hongxia Ma^{1,4}, Guangfu Jin^{1,4}, Dake Li⁵, Ni Li⁶, Zhibin Hu^{1,4,7,✉}, Dong Hang^{1,4,✉}

¹Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

²Department of Gynecologic Oncology, Affiliated Tumor Hospital of Nantong University (Nantong Tumor Hospital), Nantong, Jiangsu 226361, China;

³Department of Reproduction, the Affiliated Changzhou Maternity and Child Health Care Hospital of Nanjing Medical University, Changzhou, Jiangsu 213003, China;

⁴Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine, Nanjing Medical University, Nanjing 211166, China;

⁵Department of Gynecologic Oncology, Nanjing Maternity and Child Health Hospital, Nanjing, Jiangsu 210004, China;

⁶Program Office for Cancer Screening in Urban China, National Cancer Centre/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;

⁷State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing, Jiangsu 211166, China.

Abstract

Long noncoding RNA (lncRNA) HOTAIR and MALAT1 are implicated in the development of multiple cancers. Genetic variants within HOTAIR and MALAT1 may affect the gene expression, thereby modifying genetic susceptibility to cervical cancer. A case-control study was designed, including 1,486 cervical cancer patients and 1,536 healthy controls. Based on RegulomeDB database, 11 SNPs were selected and genotyped by using Sequenom's Mass ARRAY. Univariate and multivariate logistic regression models were used to calculate the odds ratio (OR) and 95% confidence interval (CI). We found that the A allele of rs35643724 in HOTAIR was associated with increased risk of cervical cancer, while the C allele of rs1787666 in MALAT1 was associated with decreased risk. Compared to individuals with 0–1 unfavorable allele, those with 3–4 unfavorable alleles showed 18% increased odds of having cervical cancer. Our findings suggest that HOTAIR rs35643724 and MALAT1 rs1787666 might represent potential biomarkers for cervical cancer susceptibility.

Keywords: cervical cancer, variant, long noncoding RNA, HOTAIR, MALAT1

Introduction

Cervical cancer is the fourth most common

malignancy and the fourth leading cause of cancer-related deaths in global women^[1]. The annual estimation of newly diagnosed cervical cancer reaches

✉Corresponding authors: Dong Hang and Zhibin Hu, Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, No. 101 Longmian Ave, Nanjing, Jiangsu 211166, China. Tel/Fax: +86-25-8686-8438/+86-25-8686-8437, E-mails: hangdong@njmu.edu.cn and zhibin_hu@njmu.edu.cn.

Received 01 October 2018, Revised 03 December 2019, Accepted 26 February 2019, Epub 30 April 2019

CLC number: R737.33, Document code: A

The authors reported no conflict of interests.

This is an open access article under the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited.

approximately 527,600, and more than 85% occur in the developing countries^[2]. Cervical infection with high-risk human papillomavirus (HPV) has been recognized as the major risk factor for the disease^[3]. Although 80% of women are likely to acquire HPV infections during their lifetime, most of the infections regress without intervention. Less than 4% of the infections become persistent, and even fewer lead to premalignant lesions and cancer^[4-5]. Previous studies demonstrated that a family history of cervical cancer was associated with cervical cancer risk, supporting the critical role of genetic susceptibility in cervical carcinogenesis^[6].

Genome-wide association studies (GWAS) that scan the entire genome for common genetic variants have identified over 450 SNPs associated with the susceptibility to different types of cancer^[7]. Of note, only 7% of these loci are located in protein-coding regions but 93% in noncoding regions^[8-9]. For cervical cancer, three GWAS analyses were undertaken, showing that several genetic variants within immune-related genes (e.g., MICA and HLA) were associated with the risk of cervical cancer^[10-12]. The identified variants only explained a small fraction of heritability of cervical cancer. Although candidate gene association studies also reported associations between cervical cancer and SNPs involved in DNA repair, cell cycle, apoptosis, and other processes, the findings were inconsistent^[13]. Therefore, further exploration of novel susceptibility biomarkers is needed to improve the identification of individuals at higher risk^[14].

Long noncoding RNAs (lncRNAs) are noncoding transcripts that are longer than 200 nucleotides and have been described as the largest subclass in the noncoding transcriptome in humans^[15]. Accumulating evidence shows that Hox transcript antisense intergenic RNA (HOTAIR) plays an oncogenic role in tumorigenicity of different cancers, such as gastric, colorectal, breast, and cervical cancer^[16]. In addition, deregulated expression of HOTAIR in HPV16-positive cervical cancer was shown to be E7-dependent, suggesting that HOTAIR could be a potential target of HPV16 E7 oncoprotein^[17]. It was also reported that genetic variations within HOTAIR might affect the interaction with HPV16 E7^[18]. Another lncRNA, termed as metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), was first found in association with lung cancer^[19]. A prior study showed that MALAT1 expression was upregulated in cervical cancer cells compared with that in normal cells, and MALAT1 might interact with HPV, promoting cell proliferation and migration^[20]. Therefore, HOTAIR and MALAT1 may serve as two

important regulators in cervical cancer development^[21].

Genetic variants within HOTAIR have been linked with the susceptibility to multiple cancers including gastric cancer^[22], breast cancer^[23], and ovarian cancer^[24]. Our previous study also suggested that functional variants within specific lncRNAs might modify the risk of cervical cancer^[25]. However, no prior study has evaluated the association of functional variants of HOTAIR and MALAT1 with risk of cervical cancer.

Therefore, to extend our knowledge, we screened potentially functional SNPs within HOTAIR and MALAT1 based on RegulomeDB database which includes high-throughput, experimental evidence for genetic variants^[26]; and genotyped 11 SNPs with a case-control study of 1,486 cervical cancer cases and 1,536 age-matched healthy controls from a Han Chinese population.

Materials and methods

Study population

This study was approved by the institutional review board of Nanjing Medical University and all participants provided written informed consent before enrollment. The criteria for participants' enrollment were described previously^[25]. In brief, a total of 1,486 newly diagnosed and histologically confirmed cervical cancer patients were consecutively recruited from the First Affiliated Hospital of Nanjing Medical University and the Nantong Tumor Hospital in Jiangsu Province from March 2006 to December 2010. All cases were histologically confirmed and those having a history of cancer or having metastasized cancer from other organs were excluded. The 1,536 controls were randomly selected from a pool of more than 30,000 individuals who participated in a community-based screening program for noninfectious diseases in Changzhou, Jiangsu Province during the same period. All controls reported no history of cancer and were matched to the cases by age (± 5 years). All participants were unrelated ethnic Han Chinese and were interviewed to complete a standardized questionnaire for collecting information on demographic data, menstrual and reproductive history, and environmental exposure history such as smoking and alcohol drinking. After the interview, approximately 5 mL of venous blood sample was collected from each participant.

SNP selection

We used RegulomeDB (<http://regulome.stanford.edu>), which includes high-throughput, experimental data sets from the Encyclopedia of DNA Elements

(ENCODE) and other sources, to identify putative functional variants. RegulomeDB presents a scoring system with categories ranging from 1 to 6 based on the degree of the experimental and computational functional consequence of a given variant, and the lower score indicates the stronger evidence for a variant to be located in a functional region. We selected the SNPs ranging from 20 kb upstream and downstream of HOTAIR and MALAT1, and included the SNPs with RegulomeDB scores ranging from 1 to 3a. Based on the criteria of minor allele frequency (MAF) >0.05 and linkage disequilibrium (LD) <0.8 in Han Chinese, we found a total of 14 potentially functional SNPs.

Genotyping and quality control

Genomic DNA was isolated from leukocyte pellets of venous blood by standard phenol-chloroform method. Candidate SNPs were genotyped by using Sequenom's Mass ARRAY® iPLEX assay according to the manufacturer's instructions. Genotyping was performed blindly without knowing the status of cases and controls. Due to the failure of primer design, three SNPs were excluded and 11 SNPs were successfully genotyped with a call rate >90%. Two water controls in each 384-well plate were used as blank controls for quality control and approximately 20% of samples were randomly selected to be repeated, yielding a concordance rate of >99%. The samples with overall genotype completion rates <90% were excluded, leaving 1,356 cervical cancer cases and 1,496 healthy controls in the final analysis.

Statistical analysis

The difference in demographic characteristics between cases and controls was evaluated by χ^2 test for categorical variables and Student's *t*-test for continuous variables. Deviation of genotype distribution from the Hardy-Weinberg equilibrium for each SNP was tested by a goodness-of-fit χ^2 test among controls. The associations between SNPs and cervical cancer risk were estimated by odds ratios (OR) and 95% confidence intervals (CI) from univariate and multivariate logistic regression analyses in three genetic models (additive, dominant, and recessive). Each model makes different assumptions about the genetic effect, as described elsewhere^[27]. The χ^2 -based Q test was used to assess the heterogeneity between subgroups. The statistical analyses were performed using SPSS 17.0 and PLINK software. All tests were two-sided and $P<0.05$ was defined as statistically significant.

Results

Table 1 shows the characteristics of 1,356 cervical cancer cases and 1,496 healthy controls. We found that cases had lower menarche age ($P<0.01$), higher parity ($P<0.01$), and higher proportion of smoking ($P<0.001$) than controls. No statistically significant difference was observed in the distribution of age, menopausal status, and family history of cervical cancer between cases and controls.

Among 11 genetic variants, rs1194337 was excluded as it was not in Hardy-Weinberg equilibrium among the controls ($P<0.01$). As shown in **Table 2**, by using multivariate logistic regression, we found that the A allele of rs35643724 in HOTAIR was significantly associated with increased risk of cervical cancer (additive model: adjusted OR=1.13, 95% CI=1–1.28, $P=0.047$). In **Table 3**, the C allele of rs1787666 in MALAT1 was associated with decreased risk of cervical cancer (additive model: adjusted OR=0.86, 95% CI=0.76–0.97, $P=0.01$).

The combined effect of these two variants on cervical cancer risk was further analyzed. In **Table 4**, we found a dosage effect of rs35643724 and rs1787666 on cervical cancer risk (P for trend=0.001). Individuals with 3-4 unfavorable alleles had 18% higher odds of having cervical cancer than those with 0-1 unfavorable allele (additive model: adjusted OR=1.18, 95% CI=1.06–1.3, $P=0.002$).

We further performed stratified analysis by age, menarche age, parity, and menopausal status. As shown in **Table 5**, the association for rs35643724 was statistically significant in the subgroup of menarche age >16 (P for heterogeneity=0.01). Also, we observed that the pathogenic effects of rs35643724 A allele were prominent in the subgroups of parity 0–1, non-smokers, and no family history of cancer; the association for rs1787666 was statistically significant in the subgroups of menarche age >16, parity 0–1, premenopause, nonsmokers, and family history of cancer. However, heterogeneity tests showed nonsignificant difference between those subgroups.

Discussion

In the current study, we systematically investigated the association of potentially functional variants within lncRNA HOTAIR and MALAT1 with cervical cancer risk. We found that HOTAIR rs35643724 was associated with increased risk of cervical cancer and MALAT1 rs1787666 was associated with decreased risk, suggesting that these two variants might serve as novel susceptibility biomarkers for cervical cancer.

Previous studies have reported several HOTAIR variants (e.g., rs920778, rs7958904, and rs874945) are

Table 1 Demographic and clinical characteristics of 1,356 cervical cancer patients and 1496 healthy controls

Variables	Cases [n(%)]	Controls [n(%)]	P value
Age (years)			0.53
≤ 50	565(42.2)	614(41.0)	
>50	774(57.8)	882(59.0)	
Menarche age (years)			<0.01
≤ 16	978(73.6)	892(59.6)	
>16	351(26.4)	604(40.4)	
Parity			<0.01
0–1	556(41.6)	715(48.5)	
≥ 2	781(58.4)	759(51.5)	
Menopausal status			0.34
Premenopause	542(40.8)	584(39.0)	
Menopause	787(59.2)	912(61.0)	
Smoking history			<0.01
Never	1,279 (95.7)	1,476 (98.7)	
Ever	58 (4.3)	20 (1.3)	
Family history of cancer			0.16
No	1,090 (81.6)	1,189 (79.5)	
Yes	246 (18.4)	307 (20.5)	
Histology			
Squamous cell carcinoma	1,021(75.3)		
Adenocarcinoma	68(5.0)		
Adenosquamous carcinoma	23(1.7)		
Unknown	244(18.0)		
Stage			
CIN3	8(0.6)		
I	335(24.7)		
II	703(51.8)		
III	173(12.8)		
IV	37(2.7)		
Unknown	100(7.4)		

CIN3: cervical intraepithelial neoplasia 3.

associated with risk of different cancers, such as esophageal cancer, colorectal cancer, gastric cancer, and breast cancer^[28]. A prior case-control study including 510 cervical cancers and 713 healthy controls found an association between HOTAIR rs920778 and cervical cancer^[29], and another study of 1,209 cases and 1,348 controls identified HOTAIR rs7958904 in relation with cervical cancer^[30]. Unlike these two studies with tagging SNPs, we screened potentially functional SNPs by using RegulomeDB and provided the first evidence that HOTAIR

rs35643724 might be a causal variant associated with cervical cancer risk. We observed no strong LD between rs35643724 and those reported SNPs. On the other hand, several MALAT1 SNPs (rs1194338, rs4102217, rs619586, and rs618586) have been linked with colorectal cancer, hepatocellular cancer, and breast cancer, suggesting a critical role of MALAT1 variants in cancer susceptibility^[31–33]. We screened potentially functional variants within MALAT1 and first found that rs1787666 was associated with the susceptibility to cervical cancer.

Table 2 Association of variations within lncRNA HOTAIR and cervical cancer risk

SNP	Genotype	Cases [n(%)]	Controls [n(%)]	OR(95%CI) ^a	P value ^a	OR (95%CI) ^b	P value ^b
rs10783618	CC	757(55.83)	822(54.95)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	TC	510(37.61)	559(37.37)	0.99(0.85–1.16)	0.91	1.00(0.85–1.17)	0.99
	TT	84(5.61)	107(7.15)	0.85(0.63–1.15)	0.30	0.84(0.61–1.14)	0.26
	Dominant	-	-	0.97(0.84–1.12)	0.67	0.97(0.84–1.13)	0.72
	Recessive	-	-	0.86(0.64–1.15)	0.30	0.84(0.62–1.13)	0.25
	Additive	-	-	0.96(0.85–1.08)	0.45	0.95(0.85–1.08)	0.45
rs11170776	GG	848(62.54)	959(64.10)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	GT	391(28.83)	466(31.15)	0.95(0.81–1.12)	0.53	0.95(0.81–1.12)	0.56
	TT	52(3.83)	69(4.61)	0.85(0.59–1.24)	0.40	0.83(0.57–1.22)	0.34
	Dominant	-	-	0.94(0.80–1.10)	0.41	0.94(0.80–1.10)	0.42
	Recessive	-	-	0.87(0.60–1.25)	0.45	0.85(0.58–1.23)	0.38
	Additive	-	-	0.94(0.82–1.07)	0.34	0.93(0.82–1.07)	0.32
rs117363442	CC	1,130(83.33)	1,252(83.69)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	GC	214(15.78)	236(15.78)	1.01(0.82–1.23)	0.96	1.01(0.82–1.24)	0.95
	GG	11(0.81)	8(0.53)	1.52(0.61–3.80)	0.37	1.57(0.62–3.96)	0.34
	Dominant	-	-	1.02(0.84–1.25)	0.83	1.03(0.84–1.26)	0.81
	Recessive	-	-	1.52(0.61–3.8)	0.37	1.57(0.62–3.95)	0.34
	Additive	-	-	1.04(0.86–1.25)	0.70	1.04(0.86–1.26)	0.67
rs35643724	GG	750(55.31)	877(58.62)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	GA	518(38.20)	541(36.16)	1.12(0.96–1.31)	0.15	1.12(0.95–1.31)	0.17
	AA	86(6.34)	77(5.15)	1.31(0.95–1.80)	0.11	1.33(0.96–1.84)	0.09
	Dominant	-	-	1.14(0.99–1.33)	0.08	1.14(0.98–1.33)	0.08
	Recessive	-	-	1.25(0.91–1.71)	0.17	1.27(0.92–1.75)	0.14
	Additive	-	-	1.13(1.00–1.28)	0.048	1.13(1.00–1.28)	0.047
rs4564384	CC	1,139(84.00)	1,224(81.82)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	CT	193(14.23)	260(17.38)	0.8(0.65–0.98)	0.03	0.78(0.63–0.96)	0.02
	TT	15(1.11)	11(0.73)	1.47(0.67–3.20)	0.34	1.58(0.72–3.48)	0.26
	Dominant	-	-	0.82(0.68–1.01)	0.06	0.81(0.66–0.99)	0.04
	Recessive	-	-	1.52(0.70–3.32)	0.29	1.64(0.74–3.61)	0.22
	Additive	-	-	0.87(0.72–1.04)	0.13	0.86(0.71–1.04)	0.11
rs894739	CC	859(63.35)	995(66.51)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	TC	420(30.97)	437(29.21)	1.11(0.95–1.31)	0.20	1.14(0.96–1.34)	0.13
	TT	65(4.79)	62(4.14)	1.21(0.85–1.74)	0.29	1.18(0.82–1.7)	0.38
	Dominant	-	-	1.13(0.96–1.31)	0.13	1.14(0.98–1.34)	0.10
	Recessive	-	-	1.17(0.82–1.68)	0.38	1.13(0.79–1.63)	0.51
	Additive	-	-	1.11(0.97–1.26)	0.12	1.12(0.98–1.27)	0.11
rs920777	GG	1,232(90.86)	1,328(88.77)	1.00 (Ref)	Ref	1.00 (Ref)	Ref
	GA	119(8.78)	166(11.10)	0.77(0.60–0.99)	0.04	0.78(0.61–1.01)	0.06
	AA	5(0.37)	2(0.13)	2.70(0.52–13.92)	0.24	2.76(0.53–14.36)	0.23
	Dominant	-	-	0.80(0.62–1.02)	0.07	0.81(0.63–1.04)	0.09
	Recessive	-	-	2.77(0.54–14.27)	0.23	2.83(0.54–14.71)	0.22
	Additive	-	-	0.83(0.65–1.05)	0.12	0.84(0.66–1.07)	0.16

OR: odds ratio; CI: confidence interval; Ref: reference. Bold font presents $P < 0.05$; ^aDerived from univariate logistic regression model; ^bAdjusted for age, smoking history, menopausal status, parity and family history of cancer.

Table 3 Association of variations within lncRNA MALAT1 and cervical cancer risk

SNP	Genotype	Cases [n(%)]	Controls [n(%)]	OR(95%CI) ^a	P value ^a	OR (95%CI) ^b	P value ^b
rs4102217	GG	964(71.1)	1,108(74.1)	1.00(Ref)	Ref	1.00(Ref)	Ref
	GC	347(25.6)	360(24.1)	1.11(0.93–1.31)	0.24	1.09(0.91–1.3)	0.34
	CC	39(2.88)	28(1.87)	1.60(0.98–2.62)	0.06	1.75(1.05–2.9)	0.03
	Dominant	-	-	1.14(0.97–1.35)	0.11	1.14(0.96–1.34)	0.14
	Recessive	-	-	1.56(0.95–2.55)	0.08	1.71(1.03–2.83)	0.04
	Additive	-	-	1.16(1.00–1.34)	0.05	1.16(1.00–1.34)	0.06
rs648578	CC	1,134(83.63)	1,258(84.09)	1.00(Ref)	Ref	1.00(Ref)	Ref
	TC	208(15.34)	224(14.97)	1.03(0.84–1.27)	0.78	1.01(0.82–1.24)	0.94
	TT	11(1.03)	14(0.94)	1.11(0.53–2.34)	0.79	0.98(0.45–2.12)	0.96
	Dominant	-	-	1.04(0.85–1.26)	0.74	1.01(0.82–1.23)	0.95
	Recessive	-	-	1.10(0.52–2.33)	0.79	0.98(0.45–2.11)	0.96
	Additive	-	-	1.04(0.86–1.24)	0.71	1.00(0.83–1.21)	0.97
rs1787666	AA	778(57.37)	811(54.21)	1.00(Ref)	Ref	1.00(Ref)	Ref
	CA	489(36.06)	561(37.50)	0.91(0.78–1.06)	0.23	0.89(0.76–1.04)	0.15
	CC	78(5.75)	116(7.75)	0.70(0.52–0.95)	0.02	0.69(0.51–0.94)	0.02
	Dominant	-	-	0.87(0.75–1.01)	0.07	0.86(0.74–1.00)	0.04
	Recessive	-	-	0.73(0.54–0.98)	0.04	0.72(0.53–0.98)	0.03
	Additive	-	-	0.87(0.77–0.98)	0.02	0.86(0.76–0.97)	0.01

OR: odds ratio; CI: confidence interval; Ref: reference. Bold font presents $P < 0.05$; ^aDerived from univariate logistic regression model; ^bAdjusted for age, smoking history, menopausal status, parity and family history of cancer.

Table 4 Combined effect of rs35643724 and rs1787666 on cervical cancer risk

Unfavorable alleles carried ^a	Cases [n(%)]	Controls [n(%)]	Adjusted OR (95%CI) ^a	P value ^a
0–1	343(25.3)	457(30.5)	1.00 (Ref)	Ref
2	621(45.8)	659(44.1)	1.29(1.07–1.54)	0.01
3–4	379(27.9)	371(24.8)	1.18(1.06–1.30)	0.002
Trend	-	-	1.18(1.07–1.31)	0.001

OR: odds ratio; CI: confidence interval; Ref: reference. Bold font presents $P < 0.05$. ^aAdjusted for age, smoking history, menopausal status, parity and family history of cancer.

The RegulomeDB score of rs35643724 was 2b and ChIP-seq data suggested that the SNP was likely to affect the binding of EZH2 protein. Overexpression of EZH2 is closely related to FIGO stage, lymph node metastasis, and a poor prognosis of cervical cancer^[34–36]. Future studies are needed to confirm the functional effects of rs35643724 on the EZH2 binding and HOTAIR expression. On the other hand, the RegulomeDB score of rs1787666 was 1f and ChIP-seq data indicated that the SNP was likely to affect the

binding of multiple proteins to MALAT1. Previous studies showed that the expression of MALAT1 in cervical cancer tissues was significantly higher than in adjacent normal tissues^[37–38]. Therefore, we speculated that rs1787666 might lower the expression of MALAT1, thereby reducing the risk of cervical cancer.

This study has several limitations. First, although we performed a relatively large study, more subjects could be available for the validation. Second, although

Table 5 Stratified analyses of rs35643724 and rs1787666 associated with cervical cancer risk

Variable	rs35643724			rs1787666					
	Cases	Controls	OR(95%CI) ^a	P value ^a	P _{het} value ^b	OR(95%CI) ^a	P value ^a	P _{het} value ^b	
Age (years)					0.53			0.93	
≤50	315/212/37	352/231/31	1.08(0.89–1.30)	0.46		321/202/36	329/223/57	0.86(0.71–1.03)	0.11
>50	427/297/49	525/310/46	1.17(1.00–1.38)	0.05		447/280/42	482/338/59	0.87(0.74–1.02)	0.09
Menarche age (years)					0.01				0.11
≤16	571/354/52	522/331/39	1.04(0.89–1.22)	0.63		545/365/59	491/328/69	0.91(0.79–1.06)	0.23
>16	165/151/34	355/210/38	1.47(1.19–1.82)	<0.01		216/115/18	320/233/47	0.73(0.58–0.91)	<0.01
Parity					0.30				0.19
0–1	298/223/34	425/254/36	1.22(1.01–1.46)	0.04		320/200/33	368/276/66	0.79(0.66–0.95)	0.01
≥2	444/284/52	439/278/41	1.07(0.91–1.27)	0.39		446/283/44	426/281/49	0.93(0.79–1.10)	0.42
Menopausal status					0.59				0.46
Premenopause	304/205/32	344/209/31	1.09(0.89–1.33)	0.39		309/194/32	309/218/53	0.82(0.68–0.99)	0.04
Menopause	431/301/54	533/332/46	1.17(1.00–1.37)	0.05		451/287/45	502/343/63	0.90(0.77–1.05)	0.18
Smoke history					0.80				0.89
Ever	30/26/2	11/8/1	1.28(0.48–3.39)	0.62		28/24/4	8/10/2	0.81(0.34–1.90)	0.63
Never	712/481/84	866/533/76	1.13(1.00–1.29)	0.049		738/459/73	803/551/114	0.86(0.76–0.97)	0.02
Family history of cancer					0.65				0.11
Yes	133/99/14	178/109/20	1.07(0.81–1.42)	0.63		142/89/14	152/127/27	0.69(0.52–0.92)	0.01
No	609/407/72	699/432/57	1.15(1.00–1.32)	0.048		624/393/63	659/434/89	0.89(0.78–1.02)	0.10

OR: odds ratio; CI: confidence interval; Ref: reference. Bold font presents $P < 0.05$. ^aAdjusted for age, smoking history, menopausal status, parity and family history of cancer (except stratification variables) in additive models. ^bHeterogeneity test.

we provided functional evidence by using bioinformatics tools, functional studies need to be conducted to validate the biological effects of the two SNPs. Third, we did not collect cervical tissue specimens to detect HPV infections among the participants. It remains unknown whether the identified associations would be modified by the exposure to HPV infection.

In conclusion, our study provided that HOTAIR rs35643724 was associated with higher risk of cervical cancer, while MALAT1 rs1787666 was associated with lower risk, suggesting the role of long noncoding RNAs in cervical carcinogenesis. Further studies are warranted to confirm the biological function of the variants.

Acknowledgments

This study was supported by National Natural Science Foundation of China (81502873), the Natural Science Foundation of Jiangsu Province (BK20150997), Priority Academic Program for the Development of Jiangsu Higher Education Institutions (Public Health and Preventive Medicine), Innovation Fund of State key Laboratory of Reproductive Medicine (SKLRM-GC201802), Clinical Medicine Research Fund of the Chinese Medical Association (17020420711) and Top-notch Academic Programs Project of Jiangsu Higher Education Institutions (PPZY2015A067).

References

- [1] Torre LA, Islami F, Siegel RL, et al. Global cancer in women: burden and trends[J]. *Cancer Epidemiol Biomarkers Prev*, 2017, 26(4): 444–457.
- [2] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012[J]. *Int J Cancer*, 2015, 136(5): E359–E386.
- [3] zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application[J]. *Nat Rev Cancer*, 2002, 2(5): 342–350.
- [4] Dalstein V, Riethmuller D, Pr  tet JL, et al. Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study[J]. *Int J Cancer*, 2003, 106(3): 396–403.
- [5] Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral[J]. *JAMA*, 2002, 288(14): 1749–1757.
- [6] de Freitas AC, Gurgel APAD, Chagas BS, et al. Susceptibility to cervical cancer: an overview[J]. *Gynecol Oncol*, 2012, 126(2): 304–311.
- [7] Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives[J]. *Nat Rev Cancer*, 2017, 17(11): 692–704.
- [8] Pennisi E. Disease risk links to gene regulation[J]. *Science*, 2011, 332(6033): 1031.
- [9] Kumar V, Wijmenga C, Withoff S. From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases[J]. *Semin Immunopathol*, 2012, 34(4): 567–580.
- [10] Chen D, Juko-Pecirep I, Hammer J, et al. Genome-wide association study of susceptibility loci for cervical cancer[J]. *J Natl Cancer Inst*, 2013, 105(9): 624–633.
- [11] Shi YY, Li L, Hu ZB, et al. A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12[J]. *Nat Genet*, 2013, 45(8): 918–922.
- [12] Miura K, Mishima H, Kinoshita A, et al. Genome-wide association study of HPV-associated cervical cancer in Japanese women[J]. *J Med Virol*, 2014, 86(7): 1153–1158.
- [13] Chen D, Gyllensten U. Lessons and implications from association studies and post-GWAS analyses of cervical cancer[J]. *Trends Genet*, 2015, 31(1): 41–54.
- [14] Chen D, Enroth S, Ivansson E, et al. Pathway analysis of cervical cancer genome-wide association study highlights the MHC region and pathways involved in response to infection[J]. *Hum Mol Genet*, 2014, 23(22): 6047–6060.
- [15] Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses[J]. *Genes Dev*, 2011, 25(18): 1915–1927.
- [16] Hajjari M, Salavaty A. HOTAIR: an oncogenic long non-coding RNA in different cancers[J]. *Cancer Biol Med*, 2015, 12(1): 1–9.
- [17] Sharma S, Mandal P, Sadhukhan T, et al. Bridging links between long noncoding RNA HOTAIR and HPV oncoprotein E7 in cervical cancer pathogenesis[J]. *Sci Rep*, 2015, 5: 11724.
- [18] Sharma Saha S, Roy Chowdhury R, Mondal NR, et al. Identification of genetic variation in the lncRNA HOTAIR associated with HPV16-related cervical cancer pathogenesis[J]. *Cell Oncol (Dordr)*, 2016, 39(6): 559–572.
- [19] Gutschner T, H  mmerle M, Diederichs S. MALAT1 — a paradigm for long noncoding RNA function in cancer[J]. *J Mol Med (Berl)*, 2013, 91(7): 791–801.
- [20] Jiang Y, Li YH, Fang SJ, et al. The role of MALAT1 correlates with HPV in cervical cancer[J]. *Oncol Lett*, 2014, 7(6): 2135–2141.
- [21] Peng L, Yuan XQ, Jiang BY, et al. LncRNAs: key players and novel insights into cervical cancer[J]. *Tumour Biol*, 2016, 37(3): 2779–2788.
- [22] Pan WT, Liu LS, Wei JY, et al. A functional lncRNA HOTAIR genetic variant contributes to gastric cancer susceptibility[J]. *Mol Carcinog*, 2016, 55(1): 90–96.
- [23] Yan R, Cao JJ, Song CH, et al. Polymorphisms in lncRNA HOTAIR and susceptibility to breast cancer in a Chinese population[J]. *Cancer Epidemiol*, 2015, 39(6): 978–985.

- [24] Qiu HF, Wang XJ, Guo RX, et al. HOTAIR rs920778 polymorphism is associated with ovarian cancer susceptibility and poor prognosis in a Chinese population[J]. *Future Oncol*, 2017, 13(4): 347–355.
- [25] Han J, Zhou W, Jia MQ, et al. Expression quantitative trait loci in long non-coding RNA PAX8-AS1 are associated with decreased risk of cervical cancer[J]. *Mol Genet Genomics*, 2016, 291(4): 1743–1748.
- [26] Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB[J]. *Genome Res*, 2012, 22(9): 1790–1797.
- [27] Lettre G, Lange C, Hirschhorn JN. Genetic model testing and statistical power in population-based association studies of quantitative traits[J]. *Genet Epidemiol*, 2007, 31(4): 358–362.
- [28] Zhang ZX, Tong X, Zhang WN, et al. Association between the HOTAIR polymorphisms and cancer risk: an updated meta-analysis[J]. *Oncotarget*, 2017, 8(3): 4460–4470.
- [29] Guo LS, Lu XG, Zheng LJ, et al. Association of long non-coding RNA HOTAIR polymorphisms with cervical cancer risk in a Chinese population[J]. *PLoS One*, 2016, 11(7): e0160039.
- [30] Jin H, Lu XY, Ni J, et al. HOTAIR rs7958904 polymorphism is associated with increased cervical cancer risk in a Chinese population[J]. *Sci Rep*, 2017, 7(1): 3144.
- [31] Li YJ, Bao CZ, Gu SM, et al. Associations between novel genetic variants in the promoter region of MALAT1 and risk of colorectal cancer[J]. *Oncotarget*, 2017, 8(54): 92604–92614.
- [32] Wang BG, Xu Q, Lv Z, et al. Association of twelve polymorphisms in three onco-lncRNA genes with hepatocellular cancer risk and prognosis: a case-control study[J]. *World J Gastroenterol*, 2018, 24(23): 2482–2490.
- [33] Peng R, Luo CL, Guo QY, et al. Association analyses of genetic variants in long non-coding RNA MALAT1 with breast cancer susceptibility and mRNA expression of MALAT1 in Chinese Han population[J]. *Gene*, 2018, 642: 241–248.
- [34] Chen SQ, Zhang HM, Li JB, et al. Analyzing simultaneous positive expression of EZH2 and P53 protein to improve predictive value in cervical squamous cell carcinoma[J]. *Int J Gynecol Cancer*, 2014, 24(9): 1653–1658.
- [35] Azizmohammadi S, Azizmohammadi S, Safari A, et al. High-level expression of RIPK4 and EZH2 contributes to lymph node metastasis and predicts favorable prognosis in patients with cervical cancer[J]. *Oncol Res*, 2017, 25(4): 495–501.
- [36] Jin MF, Yang ZJ, Ye WP, et al. Prognostic significance of histone methyltransferase enhancer of zeste homolog 2 in patients with cervical squamous cell carcinoma[J]. *Oncol Lett*, 2015, 10(2): 857–862.
- [37] Yang L, Bai HS, Deng Y, et al. High MALAT1 expression predicts a poor prognosis of cervical cancer and promotes cancer cell growth and invasion[J]. *Eur Rev Med Pharmacol Sci*, 2015, 19(17): 3187–3193.
- [38] Guo FJ, Li YL, Liu Y, et al. Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion[J]. *Acta Biochim Biophys Sin (Shanghai)*, 2010, 42(3): 224–229.