IVYSPRING INTERNATIONAL PUBLISHER

Journal of Cancer

2017; 8(7): 1162-1169. doi: 10.7150/jca.18055

Research Paper

Association between SNPs in Long Non-coding RNAs and the Risk of Female Breast Cancer in a Chinese Population

Tao Xu[1*], Xiu-Xiu Hu[1, 2*], Xiang-Xiang Liu[1], Han-Jin Wang[3], Kang Lin[1], Yu-Qin Pan[1], Hui-Ling Sun[1], Hong-Xin Peng[1, 2], Xiao-Xiang Chen[1, 2], Shu-Kui Wang[1]✉ and Bang-Shun He[1]✉

1. General Clinical Research Center, Nanjing First Hospital, Nanjing Medical University, Nanjing, China;
2. Medical college, Southeast University, Nanjing, China;
3. Department of General Surgery, Nanjing First Hospital, Nanjing Medical University, Nanjing, China.

* These authors contributed equally to this work.

✉ Corresponding authors: Shu-Kui Wang Email: sk_wang@njmu.edu.cn; Bang-Shun He Email: hebangshun@163.com

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (https://creativecommons.org/licenses/by-nc/4.0/). See http://ivyspring.com/terms for full terms and conditions.

Received: 2016.10.24; Accepted: 2016.12.19; Published: 2017.04.09

Abstract

Long non-coding RNAs (LncRNAs) have been reported to be involved in tumorigenesis and tumor progression. Single nucleotide polymorphisms (SNPs) in the lncRNAs also play a vital role in carcinogenesis. The aim of this study was to assess the relationships between the four selected tagSNPs (rs944289, rs3787016, rs1456315, rs7463708) in the lncRNAs and the risk of female breast cancer in a Chinese population. A case-control study was carried out involving in a total of 439 breast cancer patients and 439 age-matched healthy controls. The genotyping was performed with Sequenom MassARRAY and the expression of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (HER-2) in tumor tissues was measured by the immunohistochemistry (IHC) assay. We found that rs3787016 TT genotype (adjusted odds ratio (OR) = 1.62, 95% confidence interval (CI) = 1.09-2.41, *P* = 0.018) was associated with an increased risk of female breast cancer, especially among the patients with premenopausal status (adjusted OR = 2.55, 95% CI = 1.30-4.97, *P* = 0.006). Moreover, a statistically significant increased risk of the rs3787016 TT genotype was observed among the patients with advanced tumor stage (III and IV), poor histological grade (G3-G4), positive lymph node involvement, positive expression of ER and PR and negative expression of HER-2; rs7463708 GT and GT/GG genotype were associated with decreased risk of breast cancer in the subgroup of patients with postmenopausal status (GT versus (*vs.*) TT: adjusted OR = 0.67, 95% CI = 0.46-0.99, *P* = 0.043; GT/GG *vs.* TT: adjusted OR = 0.68, 95% CI = 0.47-0.98, *P* = 0.041) and tumor late-stage (GT *vs.* TT: adjusted OR = 0.65, 95% CI = 0.43-0.97, *P* = 0.037; GT/GG *vs.* TT: adjusted OR = 0.65, 95% CI = 0.44-0.96, *P* = 0.029). In short, rs3787016 TT genotype was associated with increased breast cancer risk and clinicopathologic features of the tumor, especially among premenopausal women.

Key words: Breast cancer; LncRNAs; SNPs

Introduction

Breast cancer (BC) is recognized as the most common malignant tumor and the leading cause of cancer-related death among females all over the world [1]. In China, breast cancer was responsible for around 268,600 new cases and 69,500 deaths in 2015 [2]. Studies have identified environmental factors as the risk for breast cancer, involving reproductive and hormonal factors including a long menstrual history, oral contraceptives use and never having children [1]. Also, genetic backgrounds play a vital role in the

etiology of breast cancer. Previous researches have reported that a number of genetic variants were associated with the risk of BC [3-5]. However, the occurrence of BC is a complex multifactorial process and the molecular mechanism remains largely unclear.

Long non-coding RNAs (LncRNAs) are a new class of regulatory non-coding RNAs with length longer than 200 nucleotides, lacking open reading frame and having no potential protein translation capacity [6]. Recently, many studies have revealed that aberrant expression of lncRNAs was significantly associated with tumorigenesis and tumor progression in different cancer types, indicating the lncRNAs act as proto-oncogene [7] or anti-oncogene [8]. In addition, lncRNAs are considered to be involved in complex pathogenesis of cancers, referring to the levels of epigenome, transcription and post-transcription [9]. LncRNAs have essential roles in multiple biological processes including chromatin remodeling, cell differentiation, cell cycle control, genome rearrangement, dosage compensation, gene imprinting and regulation of gene expression [10]. Study demonstrated that the lncRNA *HOTAIR* was overexpressed in BC tissues and participated in BC progression [9]. Also, up-regulated lncRNA *MALAT1* was detected in lung adenocarcinoma, which predicted metastasis and poor prognosis in early-stage non-small cell lung cancer [11]. Additionally, research has shown that lncRNA *MEG3* expression was down-regulated in gastric cancer tissues and cell lines, and was associated with metastasis of gastric cancer by its function as a competing endogenous RNA (ceRNA) of miR-181s to regulate gastric cancer progression [12].

SNPs in lncRNAs may affect the function of target genes through altering the process of splicing and stability of mRNA conformation, leading to the modification of their interacting partners [13]. To date, the susceptibility of lncRNA SNPs to cancer risk have been investigated by numerous researches, such as lncRNA *HOTAIR* rs920778 polymorphism cause *HOTAIR* up-regulated among T allele carriers, which enhancing esophageal squamous cell carcinoma (ESCC)[14] and cervical cancer [15] risk in a Chinese population. However, Bayram *et al.* [16] reported that *HOTAIR* rs920778 CC genotype significantly increased the BC risk in a Turkish population. In addition, it was identified that lncRNA *HULC* rs7763881 may decrease the risk of HBV-related hepatocellular carcinoma in a Chinese population [17]. Subsequently, genome-wide association studies (GWAS) have identified that the C allele carriers of rs12325489 in the lincRNA-ENST00000515084 are associated with increased BC risk [18].

Based on the above backgrounds, the SNPs in the lncRNAs were the susceptibility of BC, and could be invested as biomarkers for the risk of BC. To date, there is no research to evaluate the susceptibility of lncRNA *PTCSC3* rs944289, lncRNA *POLR2E* rs3787016, lncRNA *PRNCR1* rs1456315 and lncRNA *PRNCR1* rs7463708 to BC risk. Therefore, in the current study, we selected these four tagSNPs in the lncRNAs and evaluated the relationships between the four SNPs and the risk of BC in a Chinese female population.

Materials and Methods

Study subjects

A total of 878 age-matched female subjects divided into case cohort with 439 BC patients and health cohort with 439 cancer-free individuals were enrolled in this population-based case-control study. All the patients were genetically unrelated and consecutively recruited starting from January 2008 to January 2016 in Nanjing First Hospital, Nanjing Medical University, China. Meanwhile, the health controls were randomly collected in the same hospital for their routine physical checkup at the same time period. For the cases and controls, a pretested questionnaire was used to record clinical information of each individual, such as tobacco smoking, alcohol consumption and other cancer history. Owing to less than ten individuals have the history of smoking and drinking, which may be attributed to the life style of Chinese female, and considering the very small size of participants has these two environmental factors, finally we adjusted inclusion criteria of cases as follows: (1) subjects were histologically diagnosed with primary BC; (2) with no history of smoking and drinking; (3) with no evidence of personal or family history of cancer. Selection criteria for controls included no prior history of cancer or other malignant conditions and without history of smoking and drinking. All participants have given written informed consents, and this study was approved by the Institutional Review Board of the Nanjing First Hospital.

SNPs genotyping

We collected these blood samples from each individual after their admission to the hospital. The whole blood samples of all participants collected in a test tube containing EDTA were used for genotyping assay. Genomic DNA was isolated from peripheral white blood and concentrated by using GoldMag-Mini Whole Blood Genomic DNA Purification Kit according to the manufacturer's directions (GoldMag Co. Ltd. Xian, China). The extracted DNA was stored at -80℃ until use. We

adopted the spectrometry (DU530 UV/vis spectrophotometer, Beckman Instruments, Fullerton, CA, USA) to detect DNA purity. Sequenom MassARRAY Assay Design 3.0 Software was used to design Multiplexed SNP MassEXTEND assay [19]. PCR and extension primers were designed by Sequenom, Inc. Assay Design. EXO-SAP was used to digest PCR-amplified DNA, and then mixed the primer extended by IPLEX chemistry, desalted using Clean Resin (Sequenom) and spotted onto Spectrochip matrix chips. Finally, results were detected by Mass Spectrometer. All samples were genotyped by Sequenom MassARRAY RS1000 according to the manufacturer's protocol. The final data was managed and analysed by Sequenom Typer 4.0 Software [19, 20].

Immunohistochemistry (IHC) assay

The immunohistochemistry (IHC) assay was applied to evaluate the expression of estrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth factor receptor 2 (HER-2) in paraffin-embedded tumor tissue[21]. The immunohistochemical analysis was conducted following the instructions inside the kit. The monoclonal rabbit ER, PR and HER-2 antibody used in this study were purchased from Spring Bioscience (Pleasanton, CA, USA).

Statistical analysis

Hardy-Weinberg equilibrium (HWE) was assessed by a goodness of fit chi-square test among the healthy controls to compare the observed genotype frequencies with the expected ones. The two-sided χ^2 test and independent t test were used to compare the selected variables between BC patients and healthy controls. Associations of the four SNPs with BC risks were estimated by a logistic regression model with odds ratios (ORs) and 95% confidence intervals (CIs), corresponding *p* values after adjustment for age and menopausal status. All statistical analysis was performed by using SPSS 23.0 for Windows (SPSS, Chicago, IL) and the P value < 0.05 was considered to be statistically significant.

Results

A total of 878 age-matched Chinese women subjects (439 BC patients and 439 healthy controls) were enrolled in this population-based case-control study to investigate the potential association between the four tagSNPs (rs944289, rs3787016, rs1456315, rs7463708) in the lncRNAs and BC risk. Clinicopathological features of patients with BC and healthy controls were summarized in Table 1, there were no statistically significant differences in age and menopausal status (mean age of patients *vs.* controls: 52.89±10.78 years *vs.* 52.95±10.89 years, P = 0.933; number of postmenopausal cases *vs.* controls: 241 *vs.* 229, P = 0.417). The observed genotype frequencies of the four SNPs in healthy controls were no significant deviations from the Hardy-Weinberg equilibrium (HWE) (P = 0.078 for rs944289, P = 0.144 for rs3787016, P = 0.167 for rs1456315 and P = 0.142 for rs7463708, respectively).

Table 1. Clinicopathological features of patients with breast cancer and healthy controls

Variables	Cases, n (%)	Controls, n (%)	*P* value[a]
Total participants	439	439	
Age(Mean±SD, years)	52.89±10.78	52.95±10.89	0.933
Menopausal status			0.417
Premenopausal	198(45.10%)	210(47.84%)	
Postmenopausal	241(54.90%)	229(52.16%)	
Tumor stage			
0-II	306(69.70%)		
III-IV	133(30.30%)		
Tumor grade			
G1-G2	317(72.21%)		
G3-G4	122(27.79%)		
Lymph node involvement			
Negative	211(48.06%)		
Positive	228(51.94%)		
ER			
Negative	166(37.81%)		
Positive	273(62.19%)		
PR			
Negative	205(46.70%)		
Positive	234(53.30%)		
HER-2			
Negative	92(20.96%)		
Positive	347(79.04%)		

[a]Two-sided χ^2 test and independent t test for the selected variables between cases and controls.

ER, estrogen receptor; PR, progesterone receptor; HER-2, human epidermal growth factor receptor-2

The genotypes and alleles distribution of the four tagSNPs in BC patients and healthy controls are presented in Table 2. Logistic regression analysis revealed that the rs3787016 TT homozygote (adjusted OR = 1.62, 95% CI: 1.09-2.41, P = 0.018) was associated with increased risk of BC when compared with the wild-type CC homozygote. Also, a borderline significantly increased risk was observed in the T allele of rs3787016 (adjusted OR = 1.21, 95% CI: 1.00-1.46, P = 0.052) for BC when compared with the C allele. However, no statistically significant association between rs944289, rs1456315 and rs7463708 and the risk of BC was observed among all participants, as shown in Table 2.

Table 2. Genotypes and allele frequencies of the four SNPs between patients with breast cancer and healthy controls

Genotype	Cases, n(%)	Controls, n(%)	Crude OR(95%CI)	*P* value	Adjusted OR(95%CI)[a]	*P* value[a]
rs944289						
CC	127(28.93)	115(26.20)	Reference		Reference	
CT	229(52.16)	237(53.99)	0.88(0.64,1.19)	0.400	0.88(0.64,1.20)	0.402
TT	83(18.91)	87(19.82)	0.86(0.58,1.28)	0.465	0.86(0.58,1.28)	0.453
CT/TT	312(71.07)	324(73.80)	0.87(0.65,1.17)	0.365	0.87(0.65,1.17)	0.368
Allele						
C	483(55.01)	467(53.19)	Reference		Reference	
T	395(44.99)	411(46.81)	0.93(0.77,1.12)	0.444	0.93(0.77,1.12)	0.442
rs3787016						
CC	137(31.21)	149(33.94)	Reference		Reference	
TC	209(47.61)	226(51.48)	1.01(0.75,1.36)	0.970	1.02(0.76,1.38)	0.894
TT	93(21.18)	64(14.58)	**1.58(1.07,2.34)**	**0.023**	**1.62(1.09,2.41)**	**0.018**
TC/TT	302(68.79)	290(66.06)	1.13(0.85,1.50)	0.388	1.13(0.85,1.51)	0.384
Allele						
C	483(55.01)	524(59.68)	Reference		Reference	
T	395(44.99)	354(40.32)	**1.21(1.01,1.46)**	**0.048**	1.21(1.00,1.46)	0.052
rs1456315						
AA	234(53.30)	244(55.58)	Reference		Reference	
GA	165(37.59)	159(36.22)	1.08(0.82,1.44)	0.584	1.08(0.81,1.43)	0.607
GG	40(9.11)	36(8.20)	1.16(0.71,1.88)	0.552	1.16(0.71,1.88)	0.555
GA/GG	205(46.70)	195(44.42)	1.10(0.84,1.43)	0.498	1.09(0.84,1.43)	0.511
Allele						
A	633(72.10)	647(73.69)	Reference		Reference	
G	245(27.90)	231(26.31)	1.08(0.88,1.34)	0.452	1.08(0.88,1.34)	0.460
rs7463708						
TT	209(47.61)	184(41.91)	Reference		Reference	
GT	190(43.28)	211(48.06)	0.79(0.60,1.05)	0.103	0.79(0.60,1.05)	0.102
GG	40(9.11)	44(10.02)	0.80(0.50,1.28)	0.355	0.81(0.50,1.30)	0.380
GT/GG	230(52.39)	255(58.09)	0.79(0.61,1.04)	0.090	0.79(0.61,1.04)	0.090
Allele						
T	608(69.25)	579(65.95)	Reference		Reference	
G	270(30.75)	299(34.05)	0.86(0.70,1.05)	0.139	0.86(0.70,1.05)	0.139

[a]Adjusted by age and menopausal status in logistic regression analysis.
The bold values indicate statistically significant data

Table 3. Stratified effects of polymorphisms in lncRNAs on breast cancer risk by menopausal status

Genotype	Premenopausal		*P* value[a]	Postmenopausal		*P* value[a]
	Patients/controls	OR(95%CI)[a]		Patients/controls	OR(95%CI)[a]	
rs944289						
CC	55/54	Reference		72/61	Reference	
CT	110/112	0.97(0.61,1.53)	0.879	119/125	0.81(0.53,1.23)	0.319
TT	33/44	0.74(0.41,1.33)	0.312	50/43	0.97(0.56,1.66)	0.903
CT/TT	143/156	0.90(0.58,1.40)	0.638	169/168	0.85(0.57,1.27)	0.428
rs3787016						
CC	61/65	Reference		76/84	Reference	
TC	97/128	0.79(0.51,1.23)	0.304	112/98	1.23(0.81,1.87)	0.324
TT	40/17	**2.55(1.30,4.97)**	**0.006**	53/47	1.25(0.75,2.08)	0.387
TC/TT	137/145	1.01(0.66,1.54)	0.970	165/145	1.24(0.85,1.82)	0.272
rs1456315						
AA	99/113	Reference		135/131	Reference	
GA	81/78	1.18(0.78,1.78)	0.431	84/81	0.99(0.67,1.47)	0.971
GG	18/19	1.08(0.54,2.17)	0.832	22/17	1.24(0.63,2.44)	0.540
GA/GG	99/97	1.16(0.79,1.72)	0.444	106/98	1.04(0.72,1.49)	0.855
rs7463708						
TT	85/87	Reference		124/97	Reference	
GT	93/100	0.96(0.63,1.44)	0.826	97/111	**0.67(0.46,0.99)**	**0.043**
GG	20/23	0.89(0.46,1.74)	0.733	20/21	0.74(0.38,1.43)	0.366
GT/GG	113/123	0.94(0.64,1.39)	0.760	117/132	**0.68(0.47,0.98)**	**0.041**

[a]Adjusted by age
The results with significant difference are in bold

Table 4. Stratified effects of SNPs in lncRNAs on breast cancer risk by the pathological characteristics of patients

Genotype	Co	Stage(0-II)		*P* value[a]	Stage(III-IV)		*P* value[a]	Grade(G1-G2)		*P* value[a]	Grade(G3-G4)		*P* value[a]	Lymph node involvement(-)		*P* value[a]	Lymph node involvement(+)		*P* value[a]
		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)	
rs944289																			
CC	115	92	Reference		35	Reference		92	Reference		35	Reference		62	Reference		65	Reference	
CT	237	159	0.84(0.60,1.18)	0.314	70	0.97(0.61,1.53)	0.879	166	0.88(0.62,1.23)	0.440	63	0.87(0.55,1.40)	0.568	110	0.86(0.59,1.26)	0.442	119	0.89(0.61,1.29)	0.536
TT	87	55	0.79(0.51,1.23)	0.300	28	1.06(0.60,1.88)	0.851	59	0.85(0.55,1.31)	0.454	24	0.91(0.50,1.64)	0.741	39	0.84(0.51,1.37)	0.478	44	0.89(0.55,1.44)	0.633
CT/TT	324	214	0.83(0.60,1.14)	0.252	98	0.99(0.64,1.54)	0.954	225	0.87(0.63,1.20)	0.393	87	0.88(0.56,1.37)	0.570	149	0.85(0.59,1.22)	0.382	163	0.89(0.62,1.27)	0.520
rs3787016																			
CC	149	103	Reference		34	Reference		101	Reference		36	Reference		70	Reference		67	Reference	
TC	226	145	0.95(0.69,1.32)	0.769	64	1.24(0.78,1.98)	0.371	151	1.00(0.72,1.39)	0.994	58	1.09(0.68,1.74)	0.716	98	0.94(0.64,1.36)	0.723	111	1.11(0.77,1.61)	0.580
TT	64	58	1.34(0.86,2.07)	0.194	35	**2.57(1.45,4.56)**	**0.001**	65	1.53(1.00,2.36)	0.053	28	**1.92(1.07,3.46)**	**0.029**	43	1.47(0.90,2.38)	0.121	50	**1.79(1.11,2.89)**	**0.016**
TC/TT	290	203	1.02(0.75,1.39)	0.911	99	1.48(0.96,2.29)	0.080	216	1.10(0.81,1.50)	0.549	86	1.23(0.79,1.90)	0.361	141	1.03(0.73,1.46)	0.858	161	1.23(0.87,1.75)	0.239
rs1456315																			
AA	244	157	Reference		77	Reference		164	Reference		70	Reference		112	Reference		122	Reference	
GA	159	119	1.16(0.85,1.59)	0.344	46	0.90(0.59,1.37)	0.620	122	1.13(0.83,1.54)	0.451	43	0.94(0.61,1.45)	0.794	77	1.05(0.74,1.50)	0.781	88	1.09(0.78,1.54)	0.610
GG	36	30	1.29(0.76,2.19)	0.340	10	0.87(0.41,1.83)	0.707	31	1.28(0.76,2.16)	0.350	9	0.86(0.39,1.88)	0.708	22	1.33(0.75,2.37)	0.335	18	1.01(0.55,1.85)	0.982
GA/GG	195	149	1.19(0.89,1.60)	0.248	56	0.90(0.61,1.33)	0.588	153	1.16(0.86,1.55)	0.330	52	0.93(0.62,1.40)	0.735	99	1.11(0.80,1.54)	0.552	106	1.08(0.78,1.49)	0.657
rs7463708																			
TT	184	139	Reference		70	Reference		148	Reference		61	Reference		98	Reference		111	Reference	
GT	211	138	0.87(0.64,1.18)	0.357	52	**0.65(0.43,0.97)**	**0.037**	138	0.81(0.59,1.09)	0.164	52	0.75(0.49,1.14)	0.172	92	0.82(0.58,1.16)	0.261	98	0.77(0.55,1.07)	0.120
GG	44	29	0.89(0.53,1.49)	0.653	11	0.65(0.32,1.33)	0.240	31	0.89(0.53,1.48)	0.645	9	0.61(0.28,1.33)	0.212	21	0.90(0.51,1.60)	0.723	19	0.72(0.40,1.31)	0.282
GT/GG	255	167	0.87(0.65,1.17)	0.344	63	**0.65(0.44,0.96)**	**0.029**	169	0.82(0.61,1.09)	0.174	61	0.72(0.48,1.08)	0.116	113	0.83(0.60,1.16)	0.276	117	0.76(0.55,1.04)	0.090

[a]Adjusted by age and menopausal status
The bold values indicate statistically significant data

Table 5. Stratified effects of SNPs in lncRNAs on breast cancer risk by the expression of ER,PR and HER-2

Genotype	Co	ER(-)		*P* value[a]	ER(+)		*P* value[a]	PR(-)		*P* value[a]	PR(+)		*P* value[a]	HER-2(-)		*P* value[a]	HER-2(+)		*P* value[a]
		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)		Ca	OR(95%CI)	
rs944289																			
CC	115	48	Reference		79	Reference		63	Reference		64	Reference		28	Reference		99	Reference	
CT	237	88	0.89(0.58,1.35)	0.572	141	0.87(0.61,1.23)	0.422	100	0.77(0.52,1.13)	0.175	129	0.97(0.67,1.42)	0.890	46	0.79(0.47,1.33)	0.367	183	0.90(0.65,1.25)	0.526
TT	87	30	0.83(0.48,1.41)	0.483	53	0.89(0.57,1.41)	0.629	42	0.87(0.54,1.41)	0.579	41	0.86(0.53,1.40)	0.544	18	0.86(0.44,1.65)	0.640	65	0.87(0.57,1.32)	0.510
CT/TT	324	118	0.87(0.58,1.29)	0.479	194	0.88(0.62,1.23)	0.437	142	0.79(0.55,1.14)	0.208	170	0.94(0.66,1.35)	0.750	64	0.80(0.49,1.31)	0.381	248	0.89(0.65,1.22)	0.472
rs3787016																			
CC	149	49	Reference		88	Reference		62	Reference		75	Reference		31	Reference		106	Reference	
TC	226	85	1.19(0.79,1.80)	0.404	124	0.93(0.66,1.32)	0.695	103	1.14(0.78,1.67)	0.495	106	0.93(0.65,1.34)	0.707	35	0.72(0.42,1.22)	0.224	174	1.11(0.81,1.53)	0.524
TT	64	32	1.53(0.90,2.63)	0.119	61	**1.69(1.08,2.64)**	**0.021**	40	1.51(0.91,2.49)	0.108	53	**1.75(1.10,2.79)**	**0.019**	26	**2.10(1.14,3.88)**	**0.017**	67	1.49(0.97,2.28)	0.069
TC/TT	290	117	1.24(0.84,1.82)	0.287	185	1.08(0.78,1.49)	0.649	143	1.20(0.84,1.72)	0.324	159	1.09(0.77,1.52)	0.638	61	1.00(0.62,1.60)	0.986	241	1.17(0.87,1.59)	0.303
rs1456315																			
AA	244	87	Reference		147	Reference		108	Reference		126	Reference		48	Reference		186	Reference	
GA	159	67	1.18(0.81,1.71)	0.403	98	1.02(0.73,1.41)	0.928	83	1.16(0.82,1.65)	0.397	82	0.99(0.71,1.40)	0.974	36	1.13(0.70,1.82)	0.622	129	1.06(0.79,1.44)	0.699
GG	36	12	0.95(0.47,1.91)	0.879	28	1.27(0.74,2.17)	0.390	14	0.89(0.46,1.72)	0.725	26	1.38(0.80,2.40)	0.247	8	1.06(0.46,2.43)	0.897	32	1.18(0.71,1.98)	0.526
GA/GG	195	79	1.14(0.80,1.63)	0.480	126	1.06(0.78,1.44)	0.696	97	1.12(0.80,1.56)	0.513	108	1.07(0.78,1.47)	0.691	44	1.12(0.71,1.76)	0.621	161	1.09(0.82,1.44)	0.575
rs7463708																			
TT	184	80	Reference		129	Reference		98	Reference		111	Reference		44	Reference		165	Reference	
GT	211	74	0.81(0.56,1.17)	0.258	116	0.78(0.57,1.08)	0.130	92	0.82(0.58,1.16)	0.251	98	0.77(0.55,1.08)	0.124	40	0.79(0.49,1.27)	0.328	150	0.79(0.59,1.07)	0.127
GG	44	12	0.64(0.32,1.27)	0.201	28	0.91(0.53,1.54)	0.711	15	0.65(0.34,1.23)	0.185	25	0.94(0.55,1.63)	0.834	8	0.71(0.31,1.63)	0.416	32	0.83(0.50,1.38)	0.479
GT/GG	255	86	0.78(0.54,1.11)	0.170	144	0.80(0.59,1.09)	0.152	107	0.79(0.56,1.10)	0.162	123	0.80(0.58,1.10)	0.162	48	0.78(0.50,1.22)	0.278	182	0.80(0.60,1.06)	0.120

[a]Adjusted by age and menopausal status
The bold values indicate statistically significant data

To identify the stratified effects of SNPs in lncRNAs on BC risk, subgroup analysis based on the menopausal status was performed and logistic regression analysis revealed that rs3787016 TT genotype carriers (adjusted OR = 2.55, 95% CI: 1.30-4.97, P = 0.006) have higher BC risk than those with wild-type CC in the premenopausal sub-cohort; in contrast, in the subgroup of postmenopausal women, rs7463708 GT genotype (adjusted OR = 0.67, 95% CI: 0.46-0.99, P = 0.043) or rs7463708 GT/GG genotype (adjusted OR = 0.68, 95% CI: 0.47-0.98, P = 0.041) was associated with decreased risk of BC when compared with the wild-type TT, as summarized in Table 3.

Moreover, we demonstrated the association of the four SNPs with the pathological characteristics (tumor stage, tumor grade and lymph node involvement) and patient's tumor tissue

characteristics (expression of ER, PR and HER-2). As shown in Table 4, we found that rs3787016 TT genotype was associated with advanced TNM (III and IV) classification (adjusted OR = 2.57, 95% CI: 1.45-4.56, *P* = 0.001), poor histological grade (G3-G4) (adjusted OR = 1.92, 95% CI: 1.07-3.46, *P* = 0.029) and positive lymph node involvement (adjusted OR = 1.79, 95% CI: 1.11-2.89, *P* = 0.016). In addition, a marginal significance of increased risk for BC with early differentiation (G1-G2) (adjusted OR = 1.53, 95% CI: 1.00-2.36, *P* = 0.053) was noticed. Also, we determined a statistically significant inverse relationship between the GT genotype (adjusted OR = 0.65, 95% CI: 0.43-0.97, *P* = 0.037) or GT/GG genotype (adjusted OR = 0.65, 95% CI: 0.44-0.96, *P* = 0.029) of rs7463708 polymorphism and tumor late-stage (III and IV). Furthermore, subgroup analysis based on expression of ER, PR and HER-2 was presented in the Table 5. Similarly, we observed that rs3787016 TT genotype was associated with increased BC risk of positive expression of ER (adjusted OR = 1.69, 95% CI: 1.08-2.64, *P* = 0.021) and PR (adjusted OR = 1.75, 95% CI: 1.10-2.79, *P* = 0.019) and negative expression of HER-2 (adjusted OR = 2.10, 95% CI: 1.14-3.88, *P* = 0.017), respectively. However, there was no significant association for the three SNPs (rs944289, rs1456315, rs7463708) in all subgroups, as shown in Table 5.

Discussion

In this population-based case-control study, we investigated the association between the four selected SNPs in the lncRNAs and the risk of female BC in a Chinese population. We observed that rs3787016 TT genotype was associated with an increased risk of female BC and clinicopathologic features of the tumor, especially among premenopausal women.

The SNP rs3787016 is in a lncRNA which located in an intron region of RNA polymerase II subunit E (*POLR2E*) gene, which encodes the fifth largest subunit of RNA polymerase II and is responsible for synthesizing messenger RNA (mRNA) in eukaryotes. Previous study suggested the functional genetic variants in lncRNA regions may contribute to carcinogenesis [22]. Moreover, the rs3787016 TT genotype was investigated to be associated with increased risk of prostate cancer in an eastern Chinese population [23], which was consistent with the result of the study containing a meta-analysis of two GWAS and a case-control study [22]; however, such a significant association could not be duplicated in a Serbian population [24]. This study indicated that the rs3787016 TT genotype was a risk factor for female BC in a Chinese population. In contrast, a case-control study of ESCC demonstrated that *POLR2E* rs3787016 CT or CT/TT genotype had a decreased risk of ESCC [25]. The different findings in the above studies might be explained as follow. Firstly, the results of association studies may vary among different cancer types. Secondly, owing to ancestral backgrounds, inter-population genetic differences including differences in allele frequencies could lead to inconsistent results. Finally, ESCC is considered to be affected by multiple environmental factors exposures and the interaction of the genetic backgrounds and environmental factors contributes to the risk of cancer, so the association of genetic variants and ESCC risk should be validated by more researches. Subsequently, subgroup analysis of this study revealed that the carriers of rs3787016 TT genotype had more evident risk effect on patients with positive expression of ER and PR. As we known, the expression of these two receptors is closely related to the menopausal status of females, and women in the premenopausal status have more estrogen and progesterone, which may be attributed to the result concluded by this study that patients with rs3787016 TT genotype have higher BC risk in the premenopausal sub-cohort.

The SNP rs1456315 is located in the prostate cancer associated noncoding RNA 1 (*PRNCR1*), which is a ~13kb lncRNA transcribed from the "gene desert" region of chromosome 8q24 (128.14-128.28Mb). SNPs in the lncRNA *PRNCR1* have been reported to influence the secondary structure of *PRNCR1* mRNA and the stability of the mRNA conformation, resulting in the occurrence and development of human diseases [26]; in addition, rs1456315 positioned in the region 2 of 8q24 was significantly associated with prostate cancer susceptibility [26]. Subsequently, study reported that rs1456315 AG genotype may contribute to a decreased risk of colorectal cancer [27]. However, in the present study, no statistically significant association was observed between the rs1456315 and the risk of BC. The inconsistent conclusions may be attributed to the different kinds of cancer. Also, this is the first study investigated the relationship between the rs1456315 and BC risk, so further large-scale studies in different populations still need to be done.

Rs7463708 overlaps with the lncRNA *PRNCR1* and is located in an enhancer of prostate cancer associated transcript 1 (*PCAT1*) 78 kb away, which is a lncRNA positioned in the 8q24 "gene desert" region and overexpressed in prostate cancer. It was reported that the *PCAT1* promoter strongly interacted with the T allele of rs7463708, suggesting that rs7463708 regulated the activation of *PCAT1* enhancer and resulted in increased *PCAT1* expression [28, 29]. In addition, *PCAT1* plays an important role in the carcinogenesis through interacting with the *GNMT*

gene involving in prostate cancer [29] and modulating mTOR signaling pathway in hepatocellular carcinoma [30], which also participating in the development of BC [31]. Thus, it is possible that the rs7463708 have potential association with BC risk. Also, our study drew a conclusion that in the sub-cohort analysis, rs7463708 GT and GT/GG genotypes were protective factors for female BC among postmenopausal status and tumor late-stage. To date, there was no study investigated the association between the rs7463708 and cancer risk except for prostate cancer risk. This is the first time investigating the association of rs7463708 and BC risk; therefore, more researches should be conducted for further study.

The SNP rs944289 at 14q13.3 is located 3.2 kb upstream of a long intergenic noncoding RNA (lincRNA) named Papillary Thyroid Carcinoma Susceptibility Candidate 3 (*PTCSC3*) and positioned in the binding site of the CCAAT/enhancer binding proteins (C/EBP) α and β [32]. The rs944289 T allele can affect binding sequence and results in missense variant and amino acid substitution (valine instead of alanine at codon 339), which may be induce multinodular goiter and papillary thyroid cancer (PTC)[33]. Up to now, all published researches were assessed the relationships between the rs944289 and differentiated thyroid carcinoma (DTC) [34-37]; however, to date, no association of other cancer types was discussed with the SNP, which may be due to the specific susceptibility of the SNP to the risk of DTC, and in this study, we also observed no significant relationships between the rs944289 and BC risk.

To our knowledge, in the present study, we investigated the association between the four selected SNPs and BC risk for the first time. Although there were some important discoveries revealed in the study, several limitations also need to be addressed. Firstly, the four selected SNPs of lncRNAs in our study may not be comprehensive because we were limited by those have been identified to have risk effect on other cancers. Also, the biological function of these lncRNAs remains largely unknown and has not been validated in experimental models, so it is difficult to explain our results. Secondly, the number of subjects in our study is not enough large and the small size in subgroup analysis may not provide statistical power to show significant results. Moreover, the clinical information of each individual is not fully reliable and detailed, which might influence the accuracy of the results. Thirdly, BC is a complex and multifactorial disease, this is not a large size population based case-control study, and the samples was not enough for the sub-group analysis, therefore, to confirm our findings, studies with more large-scale samples including different ethnic populations and detailed clinical information should be conducted. In addition, properly functional assessments also should be performed to illuminate the etiology of the BC.

In summary, this study demonstrated that rs3787016 TT genotype was associated with BC risk and clinicopathologic features of the tumor, especially among premenopausal women. Nevertheless, the results of this preliminary study need to be validated by further larger and well-designed researches.

Acknowledgments

This project was supported by grants from the National Nature Science Foundation of China (No.81472027 to S.W and No.81501820 to Y.P); Jiangsu Provincial Medical Youth Talent to B.H (QNRC2016066) and Y.P (QNRC2016074); Nanjing Medical Science and Technique Development Foundation to B.H (JQX13003) and Nanjing Medical University Science and Technique Development Foundation Project to HL. S (No.2015NJMUZD049).

Competing Interests

The authors have declared that no competing interest exists.

References

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015; 65: 87-108.
2. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA: a cancer journal for clinicians. 2016; 66: 115-32.
3. Michailidou K, Beesley J, Lindstrom S, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nature genetics. 2015; 47: 373-80.
4. Sawyer E, Roylance R, Petridis C, et al. Genetic predisposition to in situ and invasive lobular carcinoma of the breast. PLoS genetics. 2014; 10: e1004285.
5. Fanale D, Amodeo V, Corsini LR, et al. Breast cancer genome-wide association studies: there is strength in numbers. Oncogene. 2012; 31: 2121-8.
6. Mattick JS. The functional genomics of noncoding RNA. Science. 2005; 309: 1527-8.
7. Li L, Feng T, Lian Y, et al. Role of human noncoding RNAs in the control of tumorigenesis. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106: 12956-61.
8. Zhang X, Rice K, Wang Y, et al. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. Endocrinology. 2010; 151: 939-47.
9. Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010; 464: 1071-6.
10. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. Nucleic acids research. 2012; 40: 6391-400.
11. Ji P, Diederichs S, Wang W, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene. 2003; 22: 8031-41.
12. Peng W, Si S, Zhang Q, et al. Long non-coding RNA MEG3 functions as a competing endogenous RNA to regulate gastric cancer progression. Journal of experimental & clinical cancer research : CR. 2015; 34: 79.
13. Burd CE, Jeck WR, Liu Y, et al. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. PLoS genetics. 2010; 6: e1001233.
14. Zhang X, Zhou L, Fu G, et al. The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. Carcinogenesis. 2014; 35: 2062-7.
15. Guo L, Lu X, Zheng L, et al. Association of Long Non-Coding RNA HOTAIR Polymorphisms with Cervical Cancer Risk in a Chinese Population. PloS one. 2016; 11: e0160039.
16. Bayram S, Sumbul AT, Batmaci CY, et al. Effect of HOTAIR rs920778 polymorphism on breast cancer susceptibility and clinicopathologic features

in a Turkish population. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2015; 36: 3863-70.
17. Liu Y, Pan S, Liu L, et al. 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.
18. 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.
19. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. Current protocols in human genetics / editorial board, Jonathan L Haines [et al]. 2009; Chapter 2: Unit 2 12.
20. Thomas RK, Baker AC, Debiasi RM, et al. High-throughput oncogene mutation profiling in human cancer. Nature genetics. 2007; 39: 347-51.
21. Gul AE, Keser SH, Barisik NO, et al. The relationship of cerb B 2 expression with estrogen receptor and progesterone receptor and prognostic parameters in endometrial carcinomas. Diagnostic pathology. 2010; 5: 13.
22. Jin G, Sun J, Isaacs SD, et al. Human polymorphisms at long non-coding RNAs (lncRNAs) and association with prostate cancer risk. Carcinogenesis. 2011; 32: 1655-9.
23. Cao DL, Gu CY, Zhu Y, et al. Polymorphisms at long non-coding RNAs and prostate cancer risk in an eastern Chinese population. Prostate cancer and prostatic diseases. 2014; 17: 315-9.
24. Nikolic ZZ, Brajuskovic GN, Pavicevic D, et al. Assessment of possible association between rs3787016 and prostate cancer risk in Serbian population. International journal of clinical and experimental medicine. 2013; 6: 57-66.
25. Kang M, Sang Y, Gu H, et al. Long noncoding RNAs POLR2E rs3787016 C/T and HULC rs7763881 A/C polymorphisms are associated with decreased risk of esophageal cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2015; 36: 6401-8.
26. Chung S, Nakagawa H, Uemura M, et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. Cancer science. 2011; 102: 245-52.
27. Li L, Sun R, Liang Y, et al. Association between polymorphisms in long non-coding RNA PRNCR1 in 8q24 and risk of colorectal cancer. Journal of experimental & clinical cancer research : CR. 2013; 32: 104.
28. Guo H, Ahmed M, Zhang F, et al. Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. Nature genetics. 2016.
29. [No authors listed]. A Prostate Cancer Risk SNP Upregulates the Tumorigenic lncRNA PCAT1. Cancer discovery. 2016.
30. Yen CH, Lu YC, Li CH, et al. Functional characterization of glycine N-methyltransferase and its interactive protein DEPDC6/DEPTOR in hepatocellular carcinoma. Molecular medicine. 2012; 18: 286-96.
31. Wazir U, Wazir A, Khanzada ZS, et al. Current state of mTOR targeting in human breast cancer. Cancer genomics & proteomics. 2014; 11: 167-74.
32. Jendrzejewski J, He H, Radomska HS, et al. The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109: 8646-51.
33. Gao Y, Chen F, Niu S, et al. Replication and Meta-Analysis of Common Gene Mutations in TTF1 and TTF2 with Papillary Thyroid Cancer. Medicine. 2015; 94: e1246.
34. Tcheandjieu C, Lesueur F, Sanchez M, et al. Fine-mapping of two differentiated thyroid carcinoma susceptibility loci at 9q22.33 and 14q13.3 detects novel candidate functional SNPs in Europeans from metropolitan France and Melanesians from New Caledonia. International journal of cancer. 2016; 139: 617-27.
35. Pereda CM, Lesueur F, Pertesi M, et al. Common variants at the 9q22.33, 14q13.3 and ATM loci, and risk of differentiated thyroid cancer in the Cuban population. BMC genetics. 2015; 16: 22.
36. Maillard S, Damiola F, Clero E, et al. Common variants at 9q22.33, 14q13.3, and ATM loci, and risk of differentiated thyroid cancer in the French Polynesian population. PloS one. 2015; 10: e0123700.
37. Wang YL, Feng SH, Guo SC, et al. Confirmation of papillary thyroid cancer susceptibility loci identified by genome-wide association studies of chromosomes 14q13, 9q22, 2q35 and 8p12 in a Chinese population. Journal of medical genetics. 2013; 50: 689-95.