Open Access Full Text Article

ORIGINAL RESEARCH

LncRNA PRNCR1 rs1456315 and CCAT2 rs6983267 Polymorphisms on 8q24 Associated with Lung Cancer

This article was published in the following Dove Press journal:
International Journal of General Medicine

Wei-Ling Yu [1,2,]*
Jin-Jian Yao [2,3,]*
Zong-Zhou Xie [1]
Yan-Jing Huang [4]
Sha Xiao [5]

[1]Oncology Department of Haikou City People's Hospital, Haikou 570208, Hainan, People's Republic of China; [2]Key Laboratory of Emergency and Trauma, Ministry of Education of Hainan Medical University, Haikou 571199, Hainan, People's Republic of China; [3]Emergency Center of Hainan General Hospital Affiliated to Hainan Medical University, Haikou 570311, Hainan, People's Republic of China; [4]Oncology Department of Hainan General Hospital Affiliated to Hainan Medical University, Haikou 570311, Hainan, People's Republic of China; [5]School of Public Health of Hainan Medical University, Haikou 571199, Hainan, People's Republic of China

*These authors contributed equally to this work

Background: Long noncoding RNA single nucleotide polymorphisms (lncRNA-SNPs) PCAT1 rs710886, PRNCR1 rs1456315 and CCAT2 rs6983267 on 8q24 region present generalizability in the susceptibility to multiple cancers, however, the influence of rs710886, rs1456315 and rs6983267 on lung cancer has not been assessed. The aim of this study was to investigate associations between three lncRNA-SNPs and lung cancer.

Methods: A case–control study was performed on 438 patients with lung cancer and 456 healthy controls in the Han population from southern China. The collected samples were genotyped by the TaqMan genotyping, and the association with clinical characteristics, including age, gender, drinking status, smoking status, pathological types and clinical stages were analyzed. And the SNP function prediction was based on lncRNASNP2, RNAfold and GTEx.

Results: The rs1456315 T allele increased the risk of lung cancer [OR=1.95, 95% CI (1.58–2.43), *P*=0.003] compared to the rs1456315 C allele, and rs1456315 significantly increased the risk of lung cancer in the dominant model [OR=1.86, 95% CI (1.16–3.00), *P*=0.002]. The rs6983267 G allele, compared with the T allele, increased the risk of lung cancer [OR=1.29, 95% CI (1.07–1.57), *P*=0.007], and rs6983267 was identified as a risk factor for lung cancer [OR=1.28, 95% CI (1.06–1.55), *P*=0.003] in the additive model. Both rs1456315 and rs6983267 demonstrated significance after adjusting for the smoking status, drinking status and age. The structure prediction found rs6983267 and rs1456315 influence the secondary structure of its lncRNA. The results from lncRNASNP2 indicated that rs6983267 and rs1456315 change gain/loss target of miRNAs.

Conclusion: *PRNCR1* rs1456315 and *CCAT2* rs6983267 on 8q24 region are significantly associated with lung cancer in the Han population of southern China and alter the potential biological function in bioinformatic analysis, and the results further extended generalism of the susceptibility of cancer-associated lncRNA-SNPs to lung cancer and underlying mechanism involved in lung cancer.

Keywords: prostate cancer-associated transcript 1, prostate cancer noncoding RNA 1, colon cancer-associated transcript 2, long noncoding RNAs, single nucleotide polymorphism

Correspondence: Sha Xiao
School of Public Health of Hainan Medical University, Haikou 571199, Hainan, People's Republic of China
Tel +86-18808977719
Email xiaosha4226@126.com

Introduction

Lung cancer has the highest incidence and mortality of all malignant tumors, accounting for 11.6% of the total cases of cancer and 18.4% of the total cancer-related deaths.[1] Environment factors include smoking, alcohol, age, air pollution and life style have been identified to modulate risk of cancer, and the important risk factor of cigarette

smoking accounts for approximately 80% of lung cancer patients, suggesting the individual genetic factors may influence susceptibility to lung cancer. Additionally, accumulative evidence suggested that cancer-associated single nucleotide polymorphism (SNP) on chromosome 8q24 bears susceptibility in common.[2–7] Genome-wide association studies have consistently identified multiple independent regions on chromosome 8q24 associated with cancers, and SNPs in lncRNA sequence on 8q24 region also share a predisposition to cancers,[7,8] LncRNA-SNPs, SNPs in the LncRNA sequence, represent an attractive class of markers that illustrate association and modify lncRNA structure and function with a predictive biological significance. The cancer-associated lncRNA-SNP locus could alter lncRNA epigenetics in terms of allele-specific regulation and functionally modulate allele-specific interactions between miRNA: lncRNA and lncRNA secondary structural change.

It has been reported that lncRNA-SNPs on chromosome 8q24 demonstrated association with cancers,[4–7] the prostate cancer-associated transcript 1 (PCAT1) gene lncRNA-SNP rs710886 (A>G) is correlated with the risk of bladder cancer[4] and lung squamous cell carcinoma,[5] the prostate cancer noncoding RNA 1 (PRNCR1) gene lncRNA-SNP rs1456315 (T>C) is associated with prostate cancer, colorectal cancer[5] and breast cancer,[6] and the colon cancer-associated transcript 2 (CCAT2) gene lncRNA-SNP rs6983267 (T >G) is closely linked to colon cancer,[8] bladder cancer[9] and hepatocellular carcinoma.[10] However, the correlation between these cancer-associated lncRNA-SNPs (rs710886, rs1456315 and rs6983267) and lung cancer has not yet been assessed. A case–control study was performed to explore correlations between the lncRNA-SNPs and lung cancer, and analyze its potential function.

Materials and Methods
Study Population
All the participants were self-reported as unrelated Han ethnicity from Hainan Province. A total of 438 patients with lung cancer were enrolled. The enrolled patients were histologically diagnosed with lung cancer between March 2011 and July 2017 at Haikou City People's Hospital and Hainan General Hospital affiliated to Hainan Medical University. A total of 456 persons who came to the hospitals for a routine physical examination during the same time period were recruited as healthy controls, and healthy controls with a history of respiratory diseases and evidence of malignancy were excluded.

Data Collection
The characteristics of each subject, including age, sex, smoking status, drinking status, and history of diabetes, were collected via a questionnaire. The classification of smokers was as follows: never-smokers were defined as those who had never smoked in their lifetime; ever-smokers were defined as those who smoked regularly before the date of lung cancer diagnosis; and smokers were defined as those who smoked currently at the date of completion of the questionnaire. In the same manner, drinkers were classified as nondrinkers, ever-drinkers, and drinkers. The clinical features of lung cancer were collected from the patients' medical records. Clinical data were analyzed according to the eighth edition of the lung cancer stage classification.[11]

The study was approved by the Hainan Medical University Ethics Committee in accordance with the Declaration of Helsinki. Written informed consent was obtained from patients or their assigned relatives due to the patients are unconscious at emergency department or intensive care unit in affiliated hospitals of Hainan Medical University (IRB:HYLL-2019-034).

Peripheral Blood DNA Extraction and Genotyping
Genomic DNA was isolated from whole blood collected in EDTA tubes using an AxyPrep Blood Genomic DNAMiniprep Kit (Axygen Biosciences, CA, US) according to instructions provided by the manufacturer. DNA yield and purity were qualified by ultraviolet light spectroscopy for DNA analysis.

Genotyping was performed with TaqMan SNP Genotyping Assays (Applied Biosystems, California, USA) on the following SNPs (gene, SNP, ABI identifier assay mix): *PCAT1*, rs710886, C__753164210; *PRNCR1*, rs1456315, C__7531200_20; and *CCAT2*, rs6983267, C__29086771_20. The conditions used for TaqMan genotyping were as follows: 2× TaqMan genotyping master mix (ABI) 2.5 μL; 20× SNP genotyping assay mix (ABI) 0.25 μL; and template DNA 2.25 μL. The PCR conditions were 95°C for 10 min, denaturation at 95°C for 15 s, and 60°C for 90 s. A total of 40 cycles were completed.

Genotypic Function Exploration by Web-Based Bioinformatics
We used RNAfold (http://rna.tbi.univie.ac.at//cgi-bin /RNAWebSuite/RNAfold.cgi) and LncRNASNP2 (http:// bioinfo.life.hust.edu.cn/lncRNASNP/) database to predict

the biological effect of the significant SNPs on lncRNAs. RNAfold is a classic database to predict RNAs structure and energy change in RNA formation according to RNA sequence. Free energy represents the energy requires to change the secondary structure from the current RNA structure. LncRNASNP2 provides comprehensive resources of SNPs in lncRNAs, including SNP effects on lncRNA structure and lncRNA-miRNA binding interaction. And we further explored the effects of SNPs on gene expression by investigating a public database of GTEx portal (http://www.gtexportal.org/home/).

Statistical Analysis

Statistical analysis was performed using STATA 10.0 SE. The characteristics of the case and control groups were compared by the X^2 test and the Wilcoxon rank-sum test if necessary. The Hardy–Weinberg equilibrium (HWE) test was performed by the likelihood ratio X^2 test. Chi-square analysis was applied to examine differences in allelic and genotypic distribution, the Cochran–Armitage test was used for genotypic trend analysis, and genetic models fit genetic susceptibility. Multinomial logistic regression was employed to estimate the odds ratio (OR) and 95% confidence interval (CI). All P-value <0.05 was considered significant.

The false-positive report probability (FPRR) was calculated as described previously,[12] only the significant result with an FPRP value less than 0.2 was considered a noteworthy finding. The threshold set 0.2 and assigned a prior probability of 0.1 to detect an odds ratio (OR) of 0.67/1.5 (protective/risk effects) for an association. We also chose $P<0.05$ as a standard of statistical significance.

Results
Characteristics of the Participants

As listed in Table 1, the median age in the case group was older than that in the control group ($P<0.001$), while there

Table 1 Characteristics of Lung Cancer Cases and Controls

Characteristics	Cases (n=438)	Controls (n=456)	X^2/t	P
Age (y) (median (IQR))	52 (52)	48 (33)	7.85	<0.001
Age ≤50	45(49)	46(48)	0.96	0.337
Age>50	67(75)	56(57)	13.36	<0.001
Male [n (%)]	305 (69.6)	319 (70.0)	0.01	0.910
Female [n (%)]	133(30.4)	137(30.4)	0.01	0.938
Smoker [n (%)]				
Nonsmoker	201 (45.9)	316 (69.3)	13.53	<0.001
Ever-smoker	105 (24.0)	75 (16.5)	5.23	0.022
Smoker	132 (30.1)	65 (14.3)	20.99	<0.001
Drinkers [n (%)]				
Nondrinker	204 (46.6)	252 (55.3)	2.19	0.138
Ever-drinker	125 (28.5)	141 (30.9)	0.33	0.566
Drinker	109 (24.9)	63 (13.8)	11.94	<0.001
Histology [n (%)]				
Squamous cell carcinoma	156 (35.6)			
Adenocarcinoma	255 (58.32)			
Small cell carcinoma	16 (3.65)			
NSCLC other than adenocarcinoma	11 (2.5)			
Stage [n (%)]				
Ia	74 (16.89)			
Ib	13 (2.97)			
IIa	119 (27.17)			
IIb	45 (10.27)			
IIIa	86 (19.63)			
IIIb	36 (8.22)			
IVa	49 (11.19)			
IVb	16 (3.65)			

was no significant difference in sex between the two groups. Nonsmokers exhibited a decreased incidence of lung cancer, and ever-smokers, smokers, and drinkers were statistically more common in the case group than in the healthy control group.

Hardy–Weinberg (HWE) Analysis

rs710886 (98.9%), rs1456315 (99.6%) and rs6983267 (99.3%) showed successful genotyping rates. The HWE test showed no significant difference in SNPs in the control group (rs710886: X^2=0.009, P=0.927; rs1456315: X^2=3.49, P=0.062; rs6983267: X^2=0.198, P=0.657), indicating that control samples were obtained from the same population.

Associations of rs710886, rs1456315, and rs6983267 with Lung Cancer

The allelic and genotypic distributions of the genotypes and associations between the SNPs and lung cancer are summarized in Table 2. The rs1456315 T allele significantly increased the risk of lung cancer compared to the C allele, and the rs1456315 TT genotype was associated with lung cancer in both the additive and dominant models; however, there was no significant correlation between the rs1456315 genotype and pathological type or clinical stage as illustrated in Table 3. The rs6983267 G allele significantly increased the risk of lung cancer compared with the T allele. The genotypic analysis found that rs6983267 was significant in both the additive and recessive models and associated positive

Table 2 Association of rs710886, rs1456315, and rs6983267 with Lung Cancer

Genotype	Cases, n (%) (n=438)	Controls, n (%) (n=456)	X^2	OR (95% CI)	P
rs710886					
A	459 (53.3)	458 (50.4)			
G	401 (46.6)	450 (49.6)	1.46	1.12 (0.93–1.36)	0.230
AA	119 (27.7)	116 (25.5)			
AG	221 (51.4)	226 (49.8)			
GG	90 (20.9)	112 (24.7)			
Additive model (AA/AG/GG)	119/221/90	116/226/112	1.54	0.89 (0.74–1.07)	0.210
Dominant model (AA+AG vs GG)	340/90	342/112	1.75	1.24 (0.89–1.72)	0.190
Recessive model (AA vs GG+AG)	119/311	116/338	0.51	1.11 (0.82–1.52)	0.475
Homozygous and heterozygous model (AA+GG vs AG)	209/221	228/226	0.23	0.94 (0.71–1.23)	0.631
rs1456315					
T	637 (73.1)	605 (66.5)			
C	235 (26.9)	305 (33.5)	9.09	1.37 (1.11–1.68)	0.003
TT	231 (53.0)	210 (46.2)			
TC	175 (40.1)	185 (40.7)			
CC	30 (6.9)	60 (13.2)			
Additive model (TT/TC/CC)	231/175/30	210/185/60	8.76	0.74 (0.61–0.90)	0.003
Dominant model (TT+TC vs CC)	406/30	395/60	9.75	2.06 (1.27–3.37)	0.002
Recessive model (TT vs TC+CC)	231/205	210/245	4.15	1.31 (1.00–1.73)	0.042
Homozygous and heterozygous model (TT+CC vs TC)	231/175	270/185	0.53	0.94 (0.68–1.20)	0.468
rs6983267					
T	433 (49.7)	511 (56.2)			
G	439 (50.3)	401 (43.8)		0.77 (0.64–0.94)	0.007
TT	107 (24.5)	146(32.1)			
GT	219 (50.2)	219 (48.1)			
GG	110 (25.2)	90 (19.8)			
Additive model (TT/TG/GG)	107/219/110	146/219/90	7.47	1.30 (1.08–1.56)	0.006
Dominant model (TT+TG vs GG)	326/110	365/90	3.80	0.73 (0.53–1.01)	0.051
Recessive model (TT vs TG+GG)	107/329	146/309	6.24	0.69 (0.51–0.93)	0.016
Homozygous and heterozygous model (TT+GG vs TG)	217/219	236/219	0.39	0.92 (0.70–1.2)	0.531

Table 3 Clinicopathologic Characteristics of Lung Cancer Patients Stratified by Polymorphic Genotypes of rs1456315

rs1456315	TT, [n (%)]	TC, [n (%)]	CC, [n (%)]	X^2	P	RR(95% CI)
Stage						
I+II	132(52.8)	105(42.0)	13(5.2)			
III+IV	71(45.5)	70(44.9)	15(9.6)	3.42	0.065	1.35(0.98–1.86)
Tumor T status						
T1+T2	165(53.8)	126(41.0)	16(5.2)			
T3+T4	66(51.1)	49(38.0)	14(10.9)	1.57	0.210	1.23(0.89–1.71)
Lymph node status						
Negative	107(49.3)	97(44.7)	13(6.0)			
Positive	124(56.6)	78(35.6)	17(7.8)	0.87	0.352	0.87(0.64–1.03)
Cell differentiation						
Well+ Moderate	175(50.8)	143(41.6)	26(7.6)			
Poor	56(60.9)	32(34.8)	4(4.3)	3.4	0.065	0.70(0.47–1.03)

lymph node; however, there was no relationship between rs6983267 and the pathological type or clinical stage as listed in Table 4, and no significant relationship was found between the risk of lung cancer and the rs710886 genotype.

LncRNA-SNPs rs710886, rs1456315, and rs6983267 and High-Risk Factors in the Multinomial Logistic Regression Analysis

The effects of rs1456315, rs6983267, age, smoking status, drinking status, type 2 diabetes and sex on lung cancer were evaluated by multinomial logistic regression analysis. The results demonstrated that rs1456315, rs6983267, smoking status, drinking status, and age were significantly associated with lung cancer, as shown in Table 5, indicating that the

rs1456315 and rs6983267 genotypes were still associated with lung cancer after adjusting for the smoking status, drinking status, and age.

Stratification Analysis of High-Risk Factors for Lung Cancer

In the stratification analysis, we redefined ever-smokers and smokers as smokers and nonsmokers as controls, and ever-drinkers and drinkers as drinkers and nondrinkers as controls. The analysis revealed in Table 6, the rs1456315-TT significantly increased the risk of lung cancer in smokers, nondrinkers and patients older than 50 years, and the rs6983267-GG increased the risk of lung cancer in smokers and drinkers as shown in Table 7.

Table 4 Clinicopathologic Characteristics of Lung Cancer Patients Stratified by Polymorphic Genotypes of rs6983267

rs6983267	TT, [n (%)]	TG, [n (%)]	GG, [n (%)]	X^2	P	RR(95% CI)
Stage						
I+II	66(26.3)	128(51)	57(22.7)			
III+IV	41(22.2)	91(49.2)	53(28.6)	2.17	0.140	1.23(0.93–1.61)
Tumor T status						
T1+T2	79(25.7)	152(49.5)	76(24.8)			
T3+T4	28(21.7)	67(51.9)	34(26.4)	0.58	0.447	1.12(0.84–1.50)
Lymph node status						
Negative	58(26.6)	116(53.2)	44(20.2)			
Positive	49(22.5)	103(47.2)	66(22.5)	4.59	0.032	1.34(1.02–1.76)
Cell differentiation						
Well+ Moderate	88(25.5)	173(50.1)	84(24.3)			
Poor	19(20.9)	46(50.5)	26(28.6)	1.14	0.287	1.195(0.86–1.66)

Table 5 SNPs and Risk Factors in Multinomial Logistic Regression Analysis

	SD	**Z**	**OR (95% CI)**	**P**
rs1456315	0.082	2.58	0.76 (0.61–0.94)	0.010
rs6983267	0.124	2.04	1.23 (1.01–1.51)	0.042
rs710886	0.869	1.53	0.86 (0.70–1.04)	0.126
Smoking status	0.346	5.75	2.34 (1.75–3.13)	<0.001
Drinking status	0.201	2.43	1.41 (1.07–1.87)	0.015
Age	0.289	4.74	2.00 (1.51–2.66)	<0.001
Type 2 diabetes	0.152	0.30	1.06 (0.80–1.40)	0.686
Sex	0.134	1.08	0.84 (0.62–1.15)	0.281

False Positive Report Probability (FPRP) Analysis

The FPRP values for significant findings at different prior probability levels are illustrated in Table 8, as defined standard that assigned a prior probability of 0.1 to detect an odds ratio (OR) of 0.67/1.5 (protective/risk effects), the evident association for rs1456315 remained noteworthy in allelic comparison (T vs C), additive model (TT/TC/CC), dominant model (TT+TC vs CC) as well as genotypic comparison in subgroup of nondrinker between lung cancer and controls. And for rs6983267, allelic comparison (T vs G), additive model (TT/TG/GG) as well as genotypic comparison in subgroup of smokers between lung cancer and controls demonstrated significantly associations.

Function Exploration Based on Bioinformatic Database

We use LncRNASNP2 and RNAfold to predict the function of rs1456315 (T>C) and rs6983267 (T>G), LncRNASNP2 used to explored interaction between miRNA and lncRNA. The rs1456315 in LncRNA PRNCR1 (NONHSAT216393.1) may bring gain target of hsa-miR-376a-2-5p and loss target of hsa-miR-3149[Table S1 and Figure S1]. The rs6983267 in LncRNA CCAT2 (NONHSAT216396.1) might indicate gain target of hsa-miR-6820-3p, hsa-miR-627-3p, hsa-miR-5190, hsa-miR-4276, hsa-miR-3164, while loss target of hsa-miR-519e-3p, hsa-miR-371a-3p, hsa-miR-33b-3p, hsa-miR-515-3p[Table S2 and Figure S2].

In RNAfold analysis, The centroid secondary and minimum free energy (minimum free energy, MFE) structure of rs1456315 and rs6983267 are shown in Figures 1 and 2, respectively, The genotype of rs1456315 also bring changes to centroid secondary and MFE of the thermodynamic ensemble, and the MFE of the thermodynamic ensemble was alter from −23.17kcal/mol (rs1456315-A) to −27.29kcal/mol (rs1456315-G), indicating rs1456315

Table 6 Stratification Analysis of rs1456315 in the Case and Control Groups Based on the Smoking Status, Drinking Status and Age

	TT, [n (%)]	**TC, [n (%)]**	**CC, [n (%)]**	**X^2**	**P**	**OR (95% CI)**
Smokers						
Cases	133 (56.6)	89 (37.9)	13 (5.5)			
Controls	75 (47.8)	61 (38.9)	21 (13.4)	5.20	0.023	0.70 (0.51–0.95)
Nonsmokers						
Cases	98 (48.8)	86 (42.8)	17 (8.5)			
Controls	135 (45.3)	124 (41.6)	39 (13.1)	1.74	0.187	0.84 (0.64–1.09)
Drinkers						
Cases	120 (51.5)	98 (42.1)	15 (6.4)			
Controls	102 (50.0)	78 (38.2)	24 (11.8)	1.20	0.273	0.85 (0.64–1.14)
Nondrinkers						
Cases	111 (54.7)	77 (37.9)	15 (7.4)			
Controls	108 (43.0)	107 (42.6)	36 (14.3)	8.60	0.003	0.66 (0.50–0.87)
Age >50 yr						
Cases	124 (53.9)	93 (40.4)	13 (5.7)			
Controls	80 (48.2)	63 (38.0)	23 (13.9)	4.37	0.037	1.39 (1.02–1.88)
Age ≤50 yr						
Cases	107 (51.9)	82 (39.8)	17 (8.3)			
Controls	130 (45.0)	122 (42.2)	37 (12.8)	3.56	0.059	0.77 (0.53–0.98)

Note: One subject in the control group was not included in this analysis for information missing.

Table 7 Stratification Analysis of rs6983267 in the Case and Control Groups Based on the Smoking Status, Drinking Status and Age

	TT, [n (%)]	GT, [n (%)]	GG, [n (%)]	X^2	P	OR (95% CI)
Smokers						
Cases	53 (22.6)	118 (50.2)	64 (27.2)			
Controls	51 (32.5)	77 (49.0)	29 (18.5)	6.60	0.010	1.46 (1.09–1.95)
Nonsmokers						
Cases	54 (26.8)	101 (50.2)	46 (22.9)			
Controls	95 (31.9)	142 (47.7)	61 (20.5)	1.31	0.252	1.16 (0.90–1.49)
Drinkers						
Cases	59 (25.3)	111 (47.6)	63 (27.0)			
Controls	69 (33.8)	100 (49.0)	35 (17.2)	7.22	0.007	1.44 (1.10–1.88)
Nondrinkers						
Cases	48 (23.6)	108 (53.2)	47 (23.2)			
Controls	77 (30.7)	119 (47.4)	55 (21.9)	1.55	0.213	1.18 (0.91–1.54)
Age >50 y						
Cases	55 (23.9)	115 (50)	60 (26.1)			
Controls	50 (29.9)	85 (50.9)	32 (19.1)	3.29	0.070	1.30 (0.97–1.73)
Age ≤50 y						
Cases	52 (25.2)	104 (50.5)	50 (24.3)			
Controls	96 (33.3)	134 (46.5)	58 (20.1)	3.51	0.061	1.27 (0.99–1.63)

may change the structural stability of lncRNA PRNCR1. And the rs6983267 in LncRNA CCAT2 would change the secondary structure of lncRNA, moreover, the MFE of the thermodynamic ensemble was change from −45.3kcal/mol (rs6983267-G) to −46.47kcal/mol (rs6983267-T), which suggests rs6983267-T may increase the structural stability of lncRNA CCAT2, comparing to rs6983267-G.

In expression analysis for rs1456315 (T>C) and rs6983267 (T>G), we found that rs6983267 contribute to expression quantitative trait loci (eQTL) based on the public database GTEx Portal. rs6983267-GG upregulated the expression CASC8 (cancer susceptibility 8, minus strand) in whole blood (Median value GG:0.409; GT:-0.071;TT:-0.3567, *P*= 6.3e-10) [Figure S3], while rs1456315 have no found related information, we further explored rs1456315 in its haplotype, including the tag-SNPs of rs146315, rs7007694, 7841060, rs16901946, rs1016343, rs13252298, rs13254738 and rs6983561, there are no related information available.

Discussion

This study investigated the association of cancer-associated lncRNA-SNPs (rs710886, rs1456315 and rs6983267) with lung cancer. The major finding was that the lncRNA-SNP *PRNCR1* rs1456315 and lncRNA *CCAT2* rs6983267 were correlated with lung cancer in the Han population of southern China, the stratification study shows rs1456315-TT significantly increased the risk of lung cancer in smokers, nondrinkers and patients older than 50 years, and rs6983267-GG increased the risk of lung cancer in smokers and drinkers, which addressed the significance of cancer-associated lncRNA-SNPs in lung cancer. In further bioinformatic exploration, we found rs6983267 and rs1456315 may influence the secondary structure of LncRNA and affect the bind to those miRNAs, rs6983267 alter the expression in eQTL on the GTEx analysis.

The rs1456315 genotype led to increased susceptibility of lung cancer. rs1456315, which is located in the 8q24 gene desert region, exerts susceptibility to various cancers, including prostate cancer, in both Japanese[13] and Iranian[14] populations, and is related to breast cancer[6] and colorectal cancer.[15] The high-risk genotype combination (TT+TC) presented a statistically significant relation to lung cancer in our study, similar to that observed in prostate cancer. Further stratification analysis of the risk factors revealed that smokers, patients older than 50 years, and nondrinkers were at increased susceptibility to lung cancer, although age and the smoking status are commonly recognized as high-risk factors for lung cancer. Interestingly, nondrinkers

Table 8 False-Positive Report Probability (FPRP) Values for the Significant Findings

Genotype	OR(95% CI)	P	Power	Prior Probability					
				0.25	0.1	0.01	0.001	0.0001	0.00001
rs1456315(T>C)									
T vs C	1.37 (1.11–1.68)	0.003	0.81	0.009	0.027	0.234	0.755	0.969	0.997
Additive model (TT/TC/CC)	0.70 (0.61–0.9)	0.005	0.63	0.025	0.071	0.458	0.895	0.988	0.999
Dominant model (TT+TC vs CC)	2.06 (1.27–3.37)	0.004	0.45	0.026	0.075	0.467	0.898	0.989	0.999
Recessive model (TT vs TC+CC)	1.31(1.00–1.73)	0.057	0.27	0.389	0.657	0.955	0.995	1.000	1.000
Smokers	0.70 (0.51–0.95)	0.022	0.61	0.098	0.245	0.782	0.973	0.997	1.00
Nondrinkers	0.66 (0.50–0.87)	0.003	0.46	0.021	0.059	0.409	0.875	0.986	0.999
Age >50 yr	1.39 (1.02–1.88)	0.032	0.69	0.124	0.298	0.824	0.979	0.998	1.00
rs6983267(T >G)									
T vs G	0.77 (0.64–0.94)	0.010	0.91	0.032	0.092	0.526	0.918	0.991	0.999
Additive model (TT/TG/GG)	1.30 (1.08–1.56)	0.005	0.938	0.015	0.044	0.336	0.836	0.981	0.998
Recessive model (TT vs TG+GG)	0.69 (0.51–0.93)	0.014	0.58	0.072	0.188	0.718	0.963	0.996	1.000
Smokers	1.46 (1.09–1.95)	0.010	0.573	0.052	0.140	0.642	0.948	0.995	0.999
Drinkers	1.39 (1.02–1.88)	0.032	0.689	0.124	0.298	0.824	0.979	0.998	1.000

showed an increased risk of lung cancer. This result is consistent with the previous study that light drinking was associated with a decreased incidence of lung cancer,[16] therefore, the further stratification of the drinking status is required to clarify the amount needed to decrease the risk of lung cancer.

rs1456315 exerts strong linkage disequilibrium with rs7463708 and rs72725879 (r^2>0.8) based on HaploReg 4.1, indicating that the rs1456315 genotype is identical to the rs7463708 and rs72725879 genotypes, thus extending the functional significance of rs1456315. The T allele of rs1456315 can increase the gene expression of PRNCR1, which is involved in prostate carcinogenesis, possibly by changing androgen receptor affinity.[4] SNPs in 8q24 have been reported to alter the secondary structure of PRNCR1 mRNA and the stability of mRNA conformation,[17] therefore rs1456315 actively involved the process of PRNCR1 regulation might also occur in lung cancer.

rs6983267 is located at the conserved region of 8q24.21, and the lncRNA CCAT2 rs6983267-GG genotype was reported to have a reproducible association with colon cancer,[18] cervical cancer[19] and lung cancer.[20] The rs6983267 of lncRNA CCAT2 is transcribed with G (CCAT2-G) or T (CCAT2-T) in a 1.7-kb RNA transcript. The allelic difference causes changes in RNA secondary structures. Changes in secondary structures can regulate the affinity of the cleavage factor I complex, and CCAT2 alters the cleavage site, thus RNA has alternative fragments with different biological functions.[21]

rs6983267-GG increases the expression of CCAT2, and high CCAT2 expression is related to lung adenocarcinoma susceptibility and a poor response to cisplatin chemotherapy.[22] Additionally, CCAT2 rs6983267-TT genotype is associated with a decreased incidence of lung adenocarcinoma.[23] Based on the theory that T/G is a protective/risk allele, rs6983267-GG also assume a risk allele for lung adenocarcinoma. And in expression quantitative trait locus analysis of whole blood, rs6983267-GG increased expression of CASC8 (also name CCAT) in whole blood, thus, taken together, rs6983267 increase likelihood to influence expressive modulation.

The false-positive report probability (FPRP) is to assess the credibility of these positive results. After FPRP analysis, the significance showed the power that further confirmed the association of rs6983267 and rs1456315 with lung cancer.

The bioinformatic analysis indicated that both rs6983267 and rs1456315 may influence lncRNA binding to miRNAs. The rs6983267 might indicate gain target of hsa-miR-6820-3p, hsa-miR-627-3p, hsa-miR-5190, hsa-miR-4276, hsa-miR-3164, while loss target of hsa-miR-519e-3p, hsa-miR-371a-3p, hsa-miR-33b-3p, hsa-miR-515-3p. Some of them involved in tumorigenesis in current literatures, the downregulation of hsa-miR-627-3p promote osteosarcoma cell proliferation and metastasis;[24] miR-33b-3p also altered the cisplatin sensitivity of cancer cells by impairing the DNA damage response;[25] miR-371a-3 currently reported biomarker

A   Centroid secondary structure

rs1456315-A   rs1456315-G
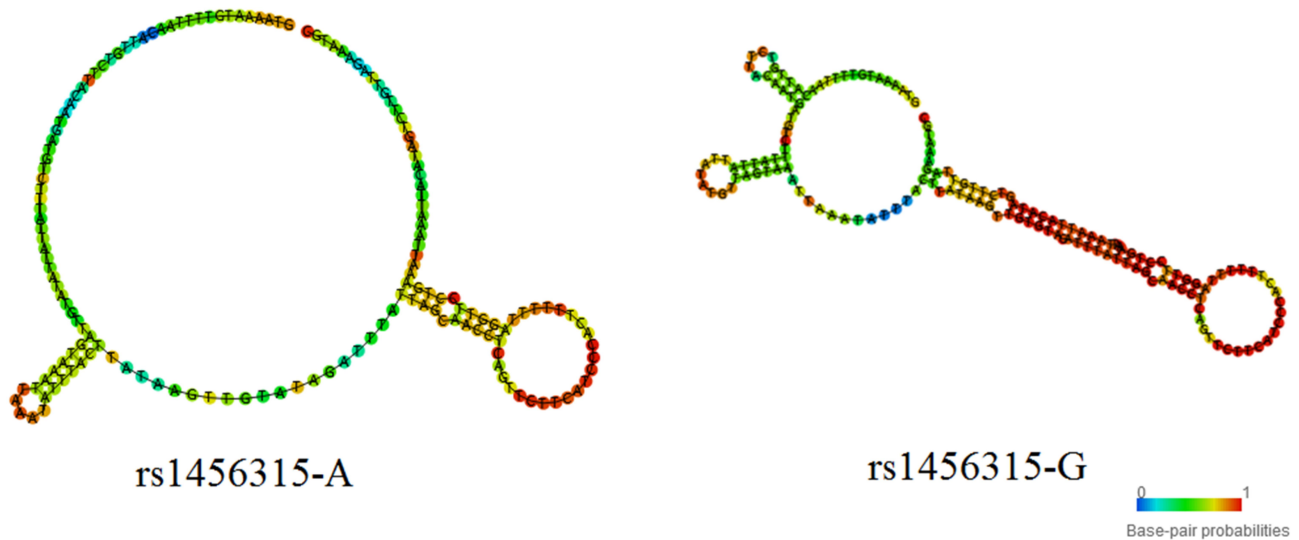
B   MFE secondary structure
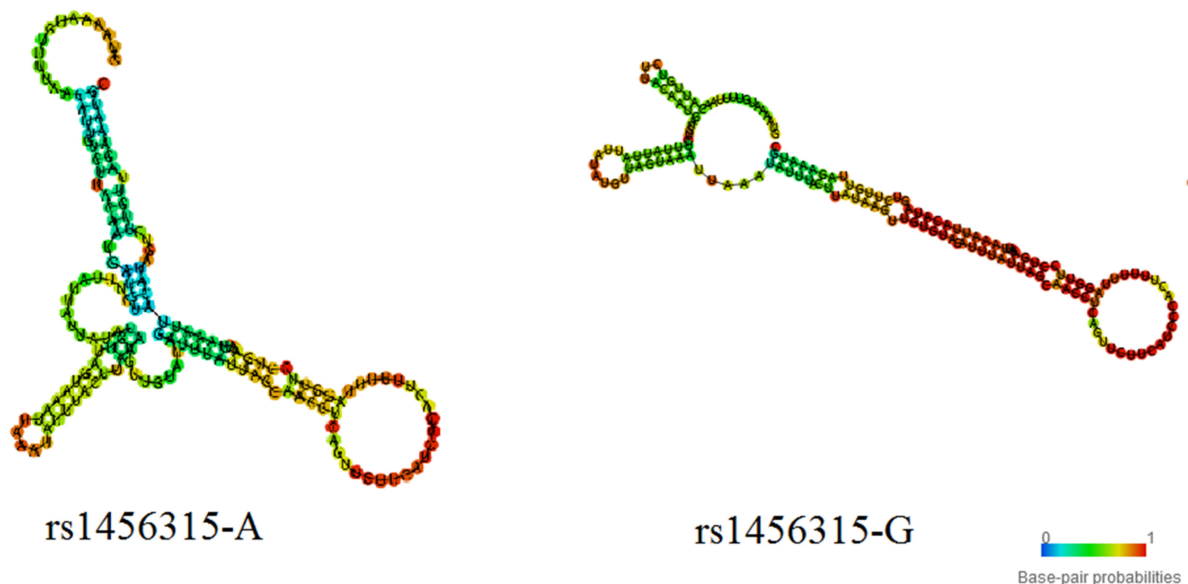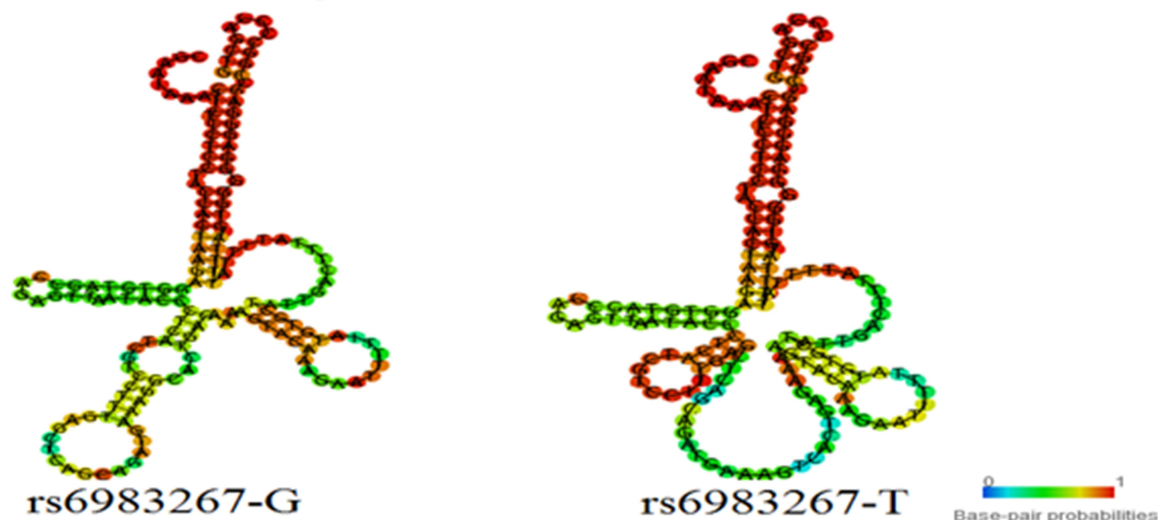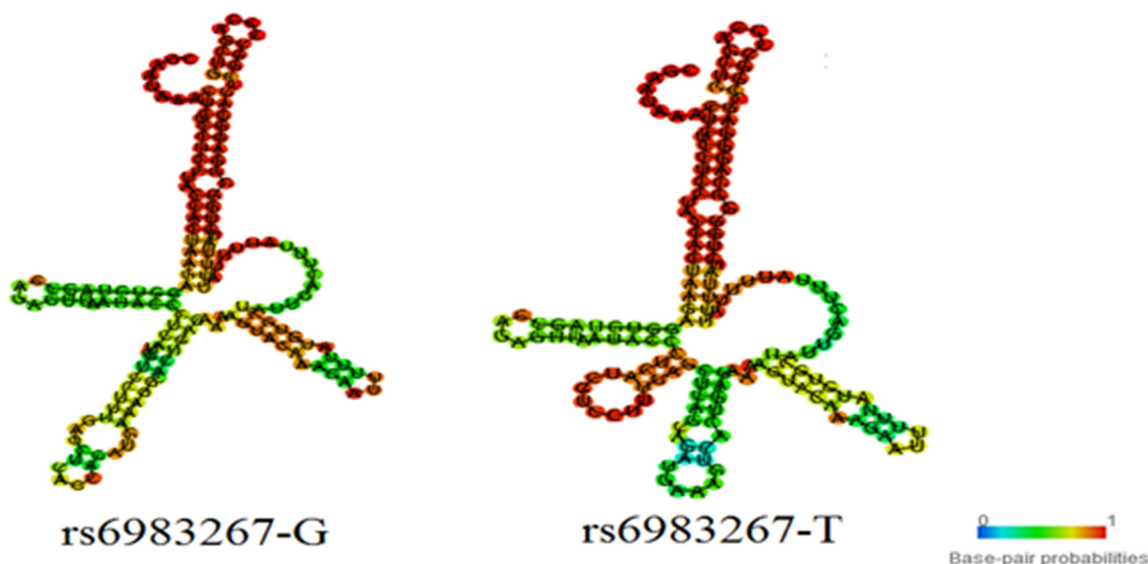

rs1456315-A   rs1456315-G

Figure 1 The RNAfold algorithm in silico predicting the genotypic impact of rs1456315 on the structure of lncRNA PRNCR1 (**A**: The prediction based on the rule of centroid secondary structure and **B**: The prediction based on the rule of MFE secondary structure).

for germ cell tumors (GCT), and clinical value of plasma miR-371a-3p level in chemotherapy naive GCT patients is a biomaker to initiate first line of chemotherapy and predict prognosis,[26,27] so the rs6983267 might regulate interaction between miR-371a-3 and lncRNA CCAT2 to alter GCT diagnosis and therapy sensitivity; miR-515-3p was markedly overexpressed in individuals with gastric carcinoma compared with that in normal gastric cells (NCs) and the surgery group ($P < 0.0001$), and yielded an increase area under the curve (AUC) value with miR-515-3p in model construction.[28] The rs1456315 may bring gain target of hsa-miR-376a-2-5p and loss target of hsa-miR-3149, and hsa-mir-3149 have reported to play important role in DNA repair and immunity by inhibiting expression of ovarian tumor protease deubiquitinase 5 (OTUD5), as a member of the ovarian tumor protease

Figure 2 The RNAfold algorithm in silico predicting the genotypic impact of rs6983267 the structure of CCAT2 (**A**: The prediction based on the rule of centroid secondary structure and **B**: The prediction based on the rule of MFE secondary structure).

family,[29] therefore rs1456315 and rs6983267 may influence binding between LncRNA and miRNA to modulate epigenetic process. Additionally, rs1456315 and rs6983267 characterize genotypic stability of LncRNA PRNCR1 and lncRNA CCAT2 secondary structure. Combined with our findings, those available evidence could shed light on the possible molecular mechanism of lung cancer.

Although we detected associations between SNP-lncRNAs and lung cancer, the identification of additional genetic risk factors would add genetic predispositions to lung cancer. There were some limitations to this study. First, prospective cohort studies are still needed to uncover the prognostic significance between SNP-lncRNAs and lung cancer incidence. Second, the study subjects were limited to the Han Chinese ethnicity, and the sample size was moderate and from two centers; therefore, large-scale studies in multiple centers across different ethnic populations are required to validate our results.

The lncRNA PRNCR1 locus rs1456315 and lncRNACCAT2 locus rs6983267 polymorphisms associated with lung cancer in the Han population of southern China, the SNP-lncRNAs may contribute to the structure and function alteration of lncRNA PRNCR1 and lncRNACCAT2, thus, the SNP-lncRNAs can be used as functional genetic markers for lung cancer and its underlying mechanism.

Acknowledgments

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors state that they have no conflicts of interest to disclose.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *Cancer J Clin*. 2015;65(2):87–108. doi:10.3322/caac.21262

2. Hu ZB, Chen JP, Tian T, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *J Clin Invest*. 2008;118 (7):2600–2608. doi:10.1172/JCI34934

3. Lin YD, Ge YQ, Wang YY, et al. The association of rs710886 in lncRNA PCAT1 with bladder cancer risk in a Chinese population. *Gene*. 2017;627(5):226–232. doi:10.1016/j.gene.2017.06.021

4. Bi YH, Cui ZG, Zhou BS, Yin ZH. Polymorphisms in long noncoding RNA-prostate cancer-associated transcript 1 are associated with lung cancer susceptibility in a Northeastern Chinese population. *DNA Cell Biol*. 2019;38(11):1357–1365. doi:10.1089/dna.2019.4834

5. Li LJ, Sun RF, Liang YD, et al. Association between polymorphisms in long non-coding RNA PRNCR1 in 8q24 and risk of colorectal cancer. *J Exp Clin Cancer Res*. 2013;32(1):104–111. doi:10.1186/ 1756-9966-32-104

6. Xu T, Hu XX, Liu XX, et al. Association between SNPs in long non-coding RNAs and the risk of female breast cancer in a Chinese population. *J Cancer*. 2017;8(7):1162–1169. doi:10.7150/jca.18055

7. Kasagi Y, Oki E, Ando K, et al. The expression of CCAT2, a novel long noncoding RNA transcript, and rs6983267 single-nucleotide polymorphism genotypes in colorectal cancers. *Oncology*. 2017;92 (1):48–54. doi:10.1159/000452143

8. Yeager M, Orr N, Hayes R, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet*. 2007;39:645–649. doi:10.1038/ng2022

9. Yadav S, Chandra A, Kumar A, Mittal B. Association of TERT-CLPTM1L and 8q24 common genetic variants with gallbladder cancer susceptibility and prognosis in North Indian population. *Biochem Genet*. 2018;56(4):267–282. doi:10.1007/s10528-018-9843-z

10. Wu ER, Hsieh MJ, Chiang WL, et al. Association of lncRNA CCAT2 and CASC8 gene polymorphisms with hepatocellular carcinoma. *Int J Environ Res Public Health*. 2019;16(16):2833–2844. doi:10.3390/ijerph16162833

11. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest*. 2017;15(1):193–203. doi:10.1016/j.chest.2016.10.010

12. Wacholder S, Chanock S, Garcia-Closas M, Ghormli LE, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*. 2004;96(6):434–442. doi:10.1093/jnci/djh075

13. Chung S, Nakagawa H, Uemura M, et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci*. 2011;102(1):245–252. doi:10.1111/j.1349-7006.2010.01737.x

14. Sattarifard H, Hashemi M, Hassanzarei S, Narouie B, Bahari G. Association between genetic polymorphisms of long non-coding RNA PRNCR1 and prostate cancer risk in a sample of the Iranian population. *Mol Clin Oncol*. 2017;7(6):1152–1158. doi:10.3892/ mco.2017.1462

15. AlMutairi M, Parine NR, Shaik JP, et al. Association between polymorphisms in PRNCR1 and risk of colorectal cancer in the Saudi population. *PLoS One*. 2019;14(9):e0220931. doi:10.1371/journal. pone.0220931

16. Choi YJ, Myung SK, Lee JH. Light alcohol drinking and risk of cancer: a meta-analysis of cohort studies. *Cancer Res Treat*. 2018;50 (2):474–487. doi:10.4143/crt.2017.094

17. Nagano T, Fraser P. No-nonsense functions for long noncoding RNAs. *Cell*. 2011;145(2):178–181. doi:10.1016/j.cell.2011.03.014

18. Ling H, Spizzo R, Atlasi Y, et al. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. *Genome Res*. 2013;23(9):1446–1461. doi:10.1101/gr.152942.112

19. Łaźniak S, Lutkowska A, Wareńczak-Florczak. Ż, et al. The association of CCAT2 rs6983267 SNP with MYC expression and progression of uterine cervical cancer in the Polish population. *Arch Gynecol Obstet*. 2018;297(5):1285–1292. doi:10.1007/s00404-018-4740-6

20. Zhang X, Chen Q, He C, et al. Polymorphisms on 8q24 are associated with lung cancer risk and survival in Han Chinese. *PLoS One*. 2012;7 (7):e41930. doi:10.1371/journal.pone.0041930

21. Redis RS, Vela LE, Lu WQ, et al. Allele-specific reprogramming of cancer metabolism by the long non-coding RNA CCAT2. *Mol Cell*. 2016;61(4):520–534. doi:10.1016/j.molcel.2016.01.015

22. Gong WJ, Yin JY, Li XP, et al. Association of well-characterized lung cancer LncRNA polymorphisms with lung cancer susceptibility and platinum-based chemotherapy response. *Tumor Biol*. 2016;37 (6):8349–8358. doi:10.1007/s13277-015-4497-5

23. Qiu M, Xu Y, Yang X, et al. CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer. *Tumor Biol*. 2014;35 (6):5375–5380. doi:10.1007/s13277-014-1700-z

24. He M, Shen P, Qiu C, Wang J. miR-627-3p inhibits osteosarcoma cell proliferation and metastasis by targeting PTN. *Aging*. 2019;11 (15):5744–5756. doi:10.18632/aging.102157

25. Xu S, Huang H, Chen Y-N, et al. DNA damage responsive miR-33b-3p promoted lung cancer cells survival and cisplatin resistance by targeting p21 WAF1/CIP1. *Cell Cycle*. 2016;15(21):2920–2930. doi:10.1080/15384101.2016.1224043

26. Spiller CM, Lobo J, Boellaard WPA, Gillis AJM, Bowles J, Looijenga LHJ. CRIPTO and miR-371a-3p are serum biomarkers of testicular germ cell tumors and are detected in seminal plasma from azoospermic males. *Cancers (Basel)*. 2020;12(3):760. doi:10.3390/cancers12030760

27. Mego M, van Agthoven T, Gronesova P, Chovanec M. Clinical utility of plasma miR-371a-3p in germ cell tumors. *J Cell Mol Med*. 2019;23(2):1128–1136. doi:10.1111/jcmm.14013

28. Han X, Xing L, Zhao H, et al. Serum miR-515-3p, a potential new RNA biomarker, is involved in gastric carcinoma. *J Cell Biochem*. 2019;120(9):15834–15843. doi:10.1002/jcb.28854

29. Bai M, Che Y, Lu K, Fu L. Analysis of deubiquitinase OTUD5 as a biomarker and therapeutic target for cervical cancer by bioinformatic analysis. *PeerJ*. 2020;8:e9146. doi:10.7717/peerj.9146