Int J Clin Exp Pathol 2017;10(6):6787-6798
www.ijcep.com /ISSN:1936-2625/IJCEP0045759

Original Article

Tag SNPs in long intergenic noncoding RNA SNHG18 contribute to susceptibility to type 2 diabetes mellitus in a southwest Chinese Han population

Hong Yi[1,5*], Rui Peng[3*], Luyu Zhang[2], Yan Sun[1,2], Huimin Peng[2,4], Handeng Liu[4], Lijuan Yu[1,2], Ailing Li[1,2], Yajuan Zhang[1,2], Wenhao Jiang[1,2], Zheng Zhang[1,2]

Departments of [1]Cell Biology and Medical Genetics, [3]Bioinformatics, Chongqing Medical University, Chongqing, China; [2]Molecular Medicine and Cancer Research Center, [4]Experimental Teaching Center, Chongqing Medical University, Chongqing, China; [5]Sichuan Science City Hospital, Sichuan Province, China. [*]Co-first authors.

Received December 4, 2016; Accepted December 25, 2016; Epub June 1, 2017; Published June 15, 2017

Abstract: By using RNA-seq, we recently identified lncRNA snhg18 which significantly down-regulated in T2DM db/db mice. In order to further study its function in human T2DM, we focused on the SNHG18, the human equivalent of snhg18. SNHG18 significantly decreased in T2DM patients, and it's a small nucleolus RNA host gene and a lincRNA. However, its potential function is still unclear. This hospital base case-control study evaluated the association between SNHG18 genetic variants and susceptibility to T2DM in a southwest Chinese Han population. We genotyped three tag SNPs (rs788321 A>T, rs2271223A>G, rs2115243A>C) of SNHG18 in 144 patients with T2DM and 88 healthy controls. Carriers of variant AT rs788321 genotype showed to increase T2DM susceptibility (*P*=0.03), while carriers of variant AC rs2115243 genotype showed to decrease T2DM susceptibility (*P*=0.005). Compared with the common genotype, AA+AT rs788321 genotype was associated with significantly increased the T2DM susceptibility and AC+CC rs2115243 genotype was associated with significantly decreased T2DM susceptibility (*P*=0.03, adjusted odds ratio [OR]=1.388, 95% confidence interval [CI]=1.032-1.868 for rs788321; *P*=0.05, adjusted OR=0.729, 95% CI=0.531-1.000 for rs2115243). Further analyses revealed that the association between T2DM susceptibility and variant genotypes of rs2115243 was more profound in non-smokers and duration of T2DM for 1~5 years. Furthermore, haplotype-based analysis revealed that individuals with A-G-C haplotype indicated a higher prevalence of T2DM (*P*=0.007, OR=4.009, 95% CI=1.347-11.929), compared to individuals with the most common A-A-A haplotype. In conclusion, our findings demonstrated that the novel lincRNA SNHG18 SNPs may contribute to susceptibility to T2DM in a southwest Chinese Han population.

Keywords: Long non-coding RNA, SNHG18, polymorphism, genotype, type 2 diabetes mellitus

Introduction

Type 2 diabetes mellitus (T2DM) is a multifactorial complex metabolic disease that resulted from the interaction of environmental factors and genetics. It was characterized by hyperglycemia, with a varying degree of insulin resistance, impaired insulin secretion and increased hepatic glucose production [1]. The number of patients with diabetes mellitus will be continually that speculated using data from the International Diabetes Federation (IDF), reaching up to 552 million by 2030, of which 90-95% will be due to T2DM [2]. In China, it is estimated that the number of patients with diabetes will increase from 20.8 million in 2000 to 42.3 million in 2030 [3]. T2DM can lead to many complications, including diabetic nephropathy, diabetic retinopathy, cardiovascular complications and diabetic neuropathy, particularly, diabetic nephropathy is a leading cause of end-stage kidney disease [4, 5]. However, the mechanism of T2DM is not fully unclear.

Various classes of long non-coding RNAs (lncRNAs) have recently been reported in many human diseases [6], including diabetes mellitus and its complications [7, 8], and emerging studies indicated that lncRNAs play key roles in various cellular contexts and diseases by different mechanisms [9-11]. The lncRNA, defined by non-protein coding transcripts >200 nucleo-

tides, is a new class of ncRNAs and has been found to be pervasively transcription in the genome, which mainly locate within the nucleus or cytosolic compartment [12]. It can be expressed from intergenic regions, antisense strands, or introns of protein coding genes, or they can be derived by alternate splicing [13]. LncRNAs can be further subcategorized into the following locus biotypes based on their location with respect to protein-coding genes: antisense lncRNAs, Intronic transcript lncRNAs, long intergenic noncoding RNAs (lincRNAs), promoter-associated lncRNAs and UTR associated lncRNAs [14]. LncRNAs have two major features, including their low expression and ability to form secondary structures that can act as DNA, RNA, and protein binding domains [12, 13]. Although the molecular mechanisms of lncRNAs still remain largely mysterious, lncRNAs may act as guides, decoys, scaffolds and signals in biological functions [15]. Some reports identified that there was some relationship between lncRNAs and T2DM [16, 17]. However, the single nucleotide polymorphisms (SNPs) of lncRNAs in T2DM remain largely undefined.

Recent studies using high-throughput RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) have identified thousands of lncRNAs [7]. By using RNA-Seq, we recently identified a number of lncRNAs that were differentially expressed in mouse models of T2DM. Of them, snhg18 (RefSeq NR_038186.1) were significantly down-regulated in db/db mice (data not shown). In order to further study the correlation to human T2DM, we focused on the SNHG18 (RefSeq NR_045196.2), the human equivalent of snhg18. LncRNA SNHG18 is a small nucleolar RNA host gene, which locus on chromosome 5p15.31 (Chr 5:9546200-9550297). The human lincRNA SNHG18 gene encodes a 4098 bp long, spliced, poly-adenylated non-coding RNA. However, to data there is no report to show the association between genetic variants of SNHG18 and the risk of any diseases, including T2DM.

In the present study, we found SNHG18 was down-expressed in T2DM patients by real-time quantitative PCR (qRT-PCR). Moreover, SNHG18 gene was evaluated by SNPs analysis in this study. Three tag SNPs (rs788321, rs2271223 and rs2115243) with a change of transcription factors were selected as candidate SNPs by bioinformatics prediction. Furthermore, we found that the genotype frequencies of rs788321 and rs2115243 polymorphisms were remarkably different in T2DM. All data indicated that SNHG18 may participate in T2DM, and two potential variants of rs788321 and rs2115243 in the promoter and intronic region of the SNHG18 gene may significantly associate with the occurrence of T2DM.

Materials and methods

Ethics statement

This study was approved by the Ethics Committee of Chongqing Medical University, and was conducted in compliance with the Helsinki Declaration. All patients involved in this study consented to participate in the study and publication of the results.

Study population

This hospital-based, case-control study was constituted of 144 patients with T2DM and 88 healthy controls who were consecutively enrolled from the first affiliated hospital of Chongqing Medical University from 2014 to 2016, and the patients who were diagnosed with T2DM for having a fasting blood-glucose ≥7 mmol/L or random blood-glucose ≥11 mmol/L or two hours blood-glucose ≥11.1 mmol/L twice according to the American Diabetes Association [18]. In addition, patients with T2DM had no diseases of cardiovascular, kidney, liver, lung and other diseases. Healthy controls were stochastically selected from Center of Health Examination during the same period. The controls had no diseases of cardiovascular, kidney, liver, lung and other diseases, which matched to the cases in terms of sex and age (±10 years). Both cases and control subjects were unconcerned, ethnically-Han Chinese individuals from southwest of china. Each patient donated 5 ml peripheral venous blood only after obtaining informed consent for the study. A structured questionnaire was used to gain basic information on the healthy controls, such as age, sex, body mass index (BMI) and smoking history. Clinical data of T2DM, including age, sex, the duration and treatment of diabetes, HbA1c and smoking habits were obtained from medical records and questionnaires. The

questionnaires were performed by experienced interviewers who were not aware of the hypothetical study. Subjects who currently or formerly smoked ≥10 cigarettes per day for at least 2 years were considered as smokers.

Total RNA isolation

The 5 ml peripheral venous blood on the empty stomach in the early morning from the patient or the healthy controls was collected and immediately added into PreAnalytiX PAXgene™ blood RNA tubes (Qiagen, Valencia, CA, USA), and total RNA was extracted by Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions.

Real-time reverse transcription (RT)-PCR analysis of lncRNA SNHG18

500 ng of total RNA was reverse transcribed into complementary DNA (cDNA) 10 μL using Primescript RT Reagent (Takara, Japan) in the Gene Amp9700 PCR system (Applied Biosystems) under the following conditions: 37°C for 15 min, 85°C for 5 s according to the manufacturer's operating instructions. QRT-PCR was used to examine lncRNA SNHG18 levels by Fast Start Universal SYBR-Green Master (Takara) using CFX96 PCR System (Bio-Rad, CA). PCR reactions were run in a 10 μl final volume containing 100 ng cDNA, 0.8 μl forward and reverse primers, 5 μl SYBR-Green and 3.2 μl ddH_2O. All intermixtures were divided into three parts. The reaction program was initiated at 95°C for 3 min, followed by 39 cycles at 95°C for 5 s, 60°C for 30 s and 72°C for 30 s, and a final elongation step of 5 s at 65°C. The primers designed by primer premier 5.0 and primer BLAST of NCBI (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome) in this study, including SNHG18 (Forward: CTCACTGGGTCCCTCTGCG, Reverse: CCCTCCCTGTTCCTCTGGTC) and GAPDH (Forward: GAAGGTGAAGGTCGGAGTC, Reverse: GAAGATGGTGATGGGATTTCC). Relative fold change of SNHG18 expression was calculated by the 2-ΔΔCT method and normalized to GAPDH.

Candidate SNP selection

The SNHG18 gene covers a 4098 bp region of chromosome 5p15.31 and includes 307 SNPs. We searched tag SNPs with the selection criterion of a minor allele frequency >5% from a genotype SNP database of the Han Chinese Beijing population (HCB data) on the HapMap Project (HapMap Rel 27, NCBI B36), which cover 8098 bp region (4098 bp SHNG18 locus and 2000 bp upstream as well as 2000 bp downstream regions of the SNHG18 gene).

Genetic analysis

The peripheral 5 ml blood on the empty stomach in the early morning was collected. Total DNA was extracted and purified from whole peripheral blood (EDTA) by using the E.Z.N.A. ® Blood DNA Kit and E.Z.N.A. ® Micro Elute® DNA Clean-Up Kit (Omega, USA), and stored at -80°C for following study. The genotypes of rs788321, rs2271223 and rs2115243, were analyzed by polymorphism-polymerase chain reaction amplification and sequencing. The primers of three tag SNPs were as follows: rs788321 (Forward: 5'-CTGCCTGGTTCTGCTGTTTG-3', Reverse: 5'-ACGGTCGTGGAGGAGTGATT-3'), rs2271223: (Forward: 5'-GACATCCTCATCCGTTCT-3', Reverse: 5'-ACTTTACTTGCTTCCCAAT-3') and rs2115243 (Forward: 5'-CTGCCACTTCCCCAGAGCC-3', Reverse: 5'-TGGGCTTCGCAGCCTCAG-3'). Amplification reactions were performed in a final volume of 25 μl containing 10× buffer solution 2.5 μl, dNTP mix (2.5 mmol/l) 2 μl, $Mgcl_2$ 2 μl, rTaq polymerase (5 U/u) 0.15 μl, PCR primers (50 pmol/l) 0.5 μl, genomic DNA (20 ng/μl) 1 μl, glycerin 1.25 μl and ddH_2O 15.1 μl. PCR program was set by 3 min at 95°C, followed by 34 cycles of 10 s at 98°C, 30 s of annealing at Tm54.9°C (rs788321)/49.4°C (rs2115243)/59°C (rs2271223) and 40 s at 72°C and a final elongation step of 5 min at 72°C. PCR products were electrophoresed with 1% agarose gels containing ethidium bromide (EB) and visualized with UV transilluminator (Bio-Rad, USA), then they were sent for sequencing (Thermo fisher scientific, USA). Genotypes were evaluated by the trained reader.

Statistical analysis

In this study, comparisons of the differences in demographic characteristics, as well as allele and genotype distributions of the SNHG18 haplotype-tag SNPs (htSNPs) (rs788321, rs2115243, and rs2271223) between T2DM patients and healthy controls were calculated using Student's t-tests and chi-square test. Associations between SNHG18 htSNPs genotypes

Table 1. Demographic information

Characteristics	Cases (n = 144)	Controls (n = 88)	*P* value
Age (years, mean ± SD)	61.73±11.47	43.56±7.70	0.213
Sex, n (%)			
Females	83 (57.64)	47 (53.41)	0.685
Males	61 (42.36)	41 (46.59)	
Smoking, n (%)			
Never	102 (70.83)	67 (76.14)	0.448
Ever	42 (29.17)	21 (23.86)	
BMI (kg/m^2)			
24	32 (22.22)	51 (57.95)	0.000
>24	112 (77.78)	37 (42.05)	
HbA1c (mmol, mean ± SD)			
~6	10 (5.65±0.25)	-	-
6.1~10	114 (11.9±1.47)	-	-
>10	20 (7.49±1.18)	-	-
Duration (years, mean ± SD)			
~1	34 (0.41±0.60)	-	-
1.1~5	58 (1.03±2.97)	-	-
5.1~10	36 (2.76±8.08)	-	-
>10	16 (5.16±16.69)	-	-
Treatment, n (%)			
Diet therapy (only)	11 (7.64)	-	-
Exercise therapy (only)	5 (3.47)	-	-
Oral hypoglycemic (only)	34 (23.61)	-	-
Insulin (only)	16 (11.11)	-	-
Combined therapy	78 (54.17)	-	-

[SD]standard deviation.

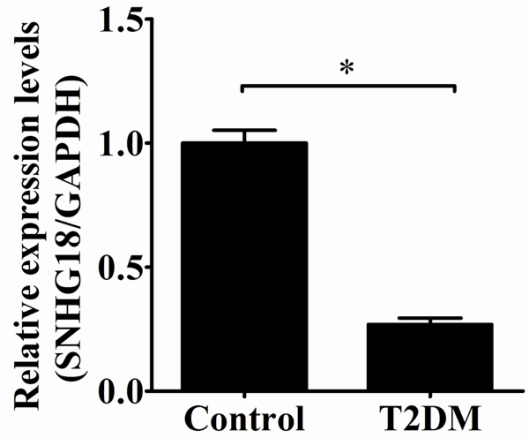

Figure 1. The relative levels of SNHG18 expression by real-time PCR in T2DM patients and controls. Each bar represents the mean ± SEM. *P<0.05 versus controls.

and risk of the development of T2DM were estimated by odds ratio (OR), which 95% confidence intervals (CIs) evaluated using unconditional logistic regression model. The crude OR was adjusted for age, sex, smoking status, hypertension and BMI. All statistical analysis were two-sided by using SPSS software (Version 20.0, SPSS Inc, Chicago, IL, USA), and a value of *P*<5% was defined as the criterion of statistical significance and presented as mean ± SEM. Linkage disequilibrium (LD) analysis and Haplotype analysis were performed by using the SHEsis software (http://analysis.bio-x.cn/myAnalysis.php) and the haploview software [19]. We only analyzed common haplotype which frequency ≥0.05 to avoid false-positive results.

Results

Characteristics of the study population

Overall, 144 patients with T2DM and 88 healthy controls were incorporated into this study. The baseline characteristics of the cases and controls were summarized in **Table 1**. There were no statistically significant differences between the cases and controls in the frequency distributions of age (*P*=0.213), sex (female 57.64 vs. 53.41, *P*= 0.685) and smoking status (non-smoker 70.83 vs. 70.61, P=0.448), indicating successful matching of subjects. However, compared with the controls, cases with T2DM had higher BMI (BMI>24 77.78 vs. 42.05, *P*<0.01).

SNHG18 expression level examination

The SNHG18 expression levels were tested by qRT-PCR in T2DM patients and controls. Data showed that SNHG18 was dramatically decreased in T2DM patients (**Figure 1**).

Bioinformatics prediction

We evaluated nine tag SNPs (rs788321, rs2271223, rs2115243, rs1651283, rs1651284, rs3798090, rs3798092, rs440527 and rs930076) which identified with the Ensembl genome (http://asia.ensembl.org/index.html),

Table 2. Selection of lncRNA SNHG18 HapMap tag SNPs with HCB data of HapMap Rel 27

Gene	SNP	Chr position	Region	MAF in CHB	Global MAF	Type of mutation	Transcription factor before mutation	Transcription factor after mutation
SNHG18	rs788321	Chr5.9544717	Promoter	0.325	0.3836	A>T	TCF7L2, FOXC2, FOXP1	GCM1, FOXP3, TFAP2A
	rs2271223	Chr5:9545934	Promoter	0.325	0.335	A>G	ZNF354C, SP1, SP2	SP8, THAP1, SOX10
	rs2115243	Chr5:9546762	Intronic	0.442	0.2885	A>C	SRY, ELF5, FOXP2	ERG, ETS1, ETV6
	rs1651283	Chr5:9547017	Intronic	0.155	0.3073	A>G	SOX10, NKX2-8	PLAG1, ZNF263, MZF1
	rs1651284	Chr5:9547161	Intronic	0.155	0.3249	A>C	THAP1, ETV3, NFATC3	SOX10, NFATC2
	rs3798090	Chr5:9541211	Promoter	0.199	0.2021	A>G	THAP1, RUNX3, ZNF354C	NFIC, DUXA, SREBF2
	rs3798092	Chr5:9541581	Promoter	0.199	0.2514	G>T	FOXP1, MEF2C, FOXP2	NFATC3, NFIC, NFATC2
	rs440527	Chr5:9548582	Intronic	0.466	0.4363	A>G	TEAD4, NKX2-8, FOXC2	RHOXF1, EMX1, MEOX2
	rs930076	Chr5:9550530	Intronic	0.218	0.2448	T>C	FOXP3, SRY, FOXG1	NFATC3, HOXC12, HOXD11

UCSC (http://genome.ucsc.edu/index.html) databases [20], dbSNP (http://www.ncbi.nlm.nih.gov/snp), Haploview version 4.2 software [21] and HapMap (http://hapmap.ncbi.nlm.nih.gov/) using the bioinformatics technology. They were located in the promoter and intronic region of the SNHG18 gene in the Chinese Han population and the minor allele frequency (MAF) was ≥0.2. We predicted binding sites between transcription factors and SNPs by TFSEARCHver1.3 software and the JASPAR database (http://jaspar.genereg.net/) [22], and we found that transcription factors were changed when before and after SNPs mutation (**Table 2**). However, the specific mechanisms are needed further studies.

Genotype of SNP sites

By polymorphism-polymerase chain reaction amplification and sequencing, we preliminary found that the SNPs of rs788321, rs2271223 and rs2115243 had statistically significant difference in T2DM, compared with the other six SNPs (rs1651283, rs1651284, rs3798090, rs3798092, rs440527 and rs930076) (data not shown), and the chromosome positions of rs788321, rs2271223 and rs2115243 were close. We further analyzed the genotypes of rs788321, rs2271223 and rs2115243 by increasing the sample size of case-control and the results were validated by sequencing (**Figure 2**).

SNHG18 variants and risk of T2DM

All genotype distributions and allele frequencies of the rs788322, rs2271223 and rs2115243 SNPs in the study population were summarized in **Table 3**. The genotype data of the selected SNPs did not deviate from Hardy-Weinberg equilibrium in the controls and cases (P=0.226 for rs788321, P=0.07 for rs2271223, and P=0.280 for rs2115243). Among the three SNPs, distributions of rs788321 and rs2115243 genotypes showed significant differences between cases and controls (P<0.05). After adjustment for the risk factors including age, sex, smoking status, BMI, the variant AT genotype of SNHG18 rs788321 were positively associated with significantly increased risk of T2DM, compared with the AA genotype (AT vs. TT: P=0.004, adjusted OR=2.602, 95% CI= 1.359-4.979). Compared with individuals with the wild-type AA genotype, subjects with the variant genotypes (AA+AT) had a significantly increased risk of T2DM (P=0.030, adjusted OR= 1.388, 95% CI=1.032-1.868). In addition, by using genetic models of additive model and dominant model, the adjusted OR for population with the variant AC and AC+CC genotypes of rs2115243 were (P=0.024, adjusted OR= 0.465, 95% CI=0.239-0.905; P=0.050, adjusted OR=0.729, 95% CI =0.531-1.000), which had a significantly decreased risk of T2DM. However, there was no significant association between cases and controls in the genotype distributions and allele frequencies of other SNHG18 polymorphism (rs2271223 A>G).

Stratified analysis of polymorphism and T2DM risk

We evaluated the effects of the variant genotypes on the risk of T2DM using stratified analyses according to age (50 years), sex, smoking status, and BMI (**Table 4**). For the SNHG18 polymorphism rs2115243, a reduced risk of T2DM associated with the variant genotypes was obvious in non-smoker subjects (P=0.035, adjusted OR=0.461, 95% CI=0.222-0.957), but not in smoker subjects (P=0.496, adjusted

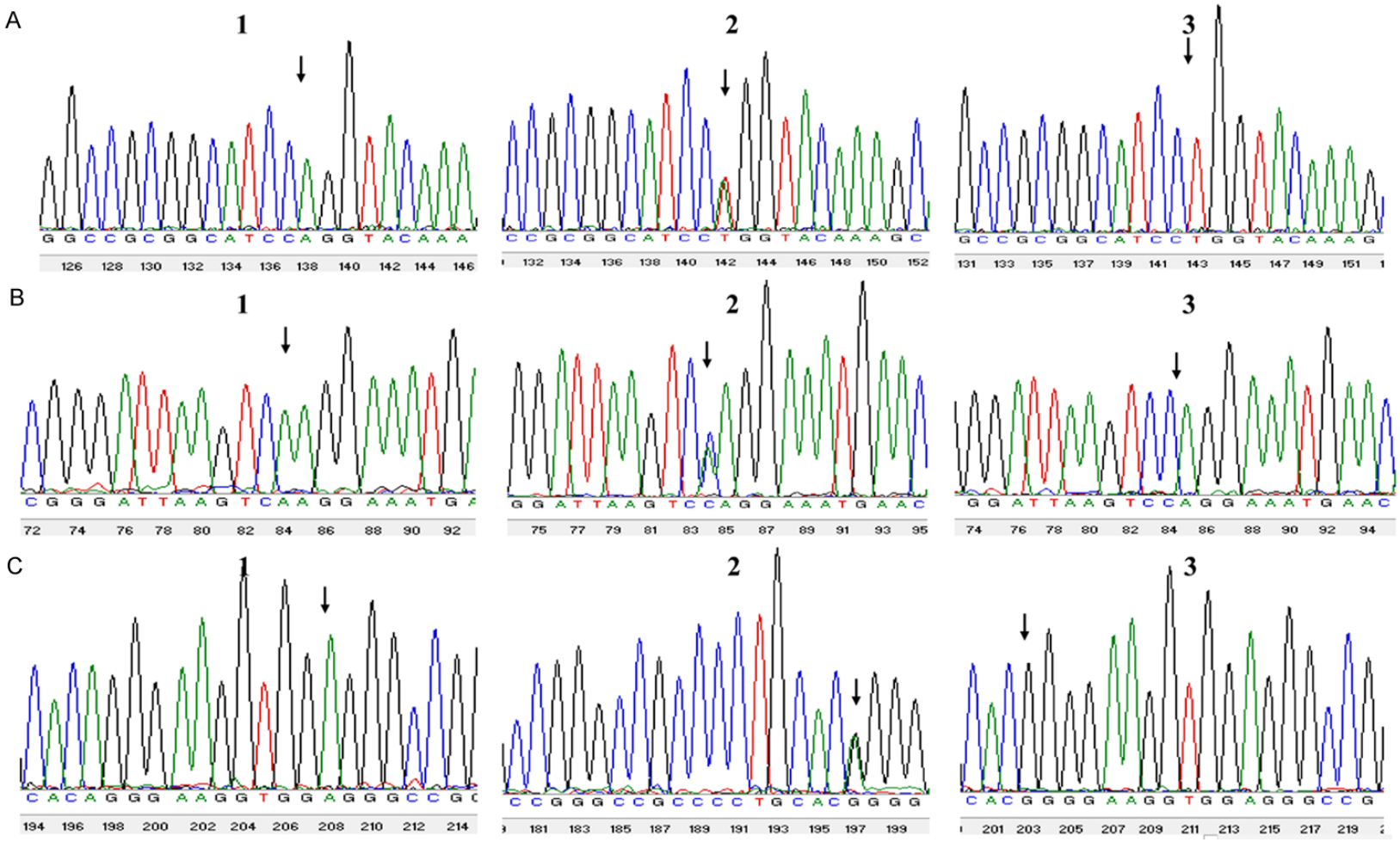

Figure 2. Sequencing data for SNHG18 genotypes. A. 1-3: rs788321-F AA, AT, TT. B. 1-3: rs2115242-F AA, AC, CC. C. 1-3: rs2271223-F AA, AG, GG.

Table 3. Genotype and allele frequencies of the SNHG18 polymorphisms among the cases and controls and their associations with risk of T2DM

SNPs	Polymorphisms	Cases (n=144) (%)	Controls (n=88) (%)	Crude OR (95% CI)	*P* value	Adjusted OR (95% CI)[1]	*P* value
rs788321	Additive model						
	TT	73 (50.69)	35 (39.77)	1		1	
	AT	50 (34.72)	45 (51.14)	**1.877 (1.062-3.318)**	**0.03**	**2.602 (1.359-4.979)**	**0.004**
	AA	21 (14.58)	8 (9.09)	1.259 (0.507-3.122)	0.619	0.880 (0.533-1.454)	0.618
	Dominant model						
	TT	73 (50.69)	35 (39.77)	1		1	
	TT+AT	123 (85.42)	80 (90.91)	1.357 (0.830-2.218)	0.223	1.558 (0.908-2.673)	0.107
	Recessive model						
	AA	21 (14.58)	8 (9.09)	1.259 (0.507-3.122)	0.619	0.880 (0.533-1.454)	0.618
	AT+AA	71 (49.31)	53 (60.23)	1.557 (0.910-2.665)	0.106	**1.388 (1.032-1.868)**	**0.03**
	Allelotype						
	T	196 (68.06)	115 (65.34)	1			
	A	92 (31.94)	61 (34.66)	1.130 (0.760-1.681)	0.546		
rs2271223	Additive model						
	AA	24 (15.97)	14 (15.91)	1		1	
	AG	53 (36.81)	32 (36.36)	1.035 (0.469-2.285)	0.932	1.014 (0.437-2.353)	0.974
	GG	67 (47.22)	42 (47.73)	1.075 (0.501-2.306)	0.853	1.081 (0.706-1.655)	0.719
	Dominant model						
	AA	24 (16.67)	14 (15.91)	1		1	
	AA+AG	77 (53.47)	46 (52.27)	1.024 (0.482-2.176)	0.951	1.012 (0.454-2.253)	0.977
	Recessive model						
	AG+GG	120 (83.33)	74 (84.09)	0.946 (0.460-1.943)	0.88	1.080 (0.494-2.362)	0.847
	GG	67 (46.53)	42 (47.73)	1.075 (0.501-2.306)	0.853	1.081 (0.706-1.655)	0.719
	Allelotype						
	A	101 (35.07)	60 (34.09)	1			
	G	187 (64.93)	116 (65.91)	1.044 (0.704-1.5490)	0.83		
rs2115243	Additive model						
	AA	58 (40.28)	21 (23.86)	1		1	
	AC	55 (38.19)	49 (55.68)	**2.461 (1.31-4.622)**	**0.005**	**0.465 (0.239-0.905)**	**0.024**
	CC	31 (23.61)	18 (20.45)	1.604 (0.746-3.449)	0.225	0.837 (0.556-1.261)	0.395
	Dominant model						
	AA	58 (40.28)	21 (23.86)	1		1	
	AA+AC	113 (78.47)	70 (79.55)	1.711 (0.957-3.060)	0.069	0.628 (0.340-1.160)	0.137
	Recessive model						
	CC	31 (21.53)	18 (20.45)	1.604 (0.746-3.449)	0.225	0.837 (0.556-1.261)	0.395
	AC+CC	86 (59.72)	67 (76.14)	**2.177 (1.203-3.939)**	**0.009**	**0.729 (0.531-1.000)**	**0.050**
	Allelotype						
	A	171 (59.38)	91 (51.70)	1			
	C	117 (40.62)	85 (48.30)	1.365 (0.936-1.992)	0.106		

The statistically significant are in the bold. [OR]odds ratio; [CI]confidence interval. [1]Adjusted for age, sex, smoking status, hypertension, BMI.

OR=1.563, 95% CI=0.430-5.682) (**Table 4**). There was no significant association between polymorphism and susceptibility to T2DM in relation to age sex, history of hypertension and BMI for rs788321 and rs2115243. Furthermore, analyses revealed that the association between T2DM susceptibility and variant genotypes of rs2115243 was more profound in duration of T2DM for 1~5 years (*P*=0.018, adjusted OR=0.320, 95% CI=0.124-0.826). There was no significant association between the variant genotypes and clinical pathological

Table 4. Stratified analyses for SNHG18 rs788321 and rs2115243 genotypes in cases and controls

Variables	(AT+AA)/TT for rs788321		Allelic odds ratios and 95% confidence intervals for rs788321		(AC+CC)/AA for rs2115243		Allelic odds ratios and 95% confidence intervals for rs2115243	
	Cases, n (%)	Control, n (%)	Adjusted OR (95% CI)[1]	*P* Value	Cases, n (%)	Control, n (%)	Adjusted OR (95% CI)[1]	*P* Value
Age (y)								
≤50	10 (6.94)/9 (6.25)	42 (47.73)/29 (32.95)	0.767 (0.277-2.122)	0.609	12 (8.33)/7 (4.86)	54 (61.36)/16 (18.18)	1.969 (0.664-5.834)	0.217
>50	62 (43.06)/63 (43.75)	10 (11.36)/7 (7.95)	0.689 (0.247-1.925)	0.475	79 (54.86)/46 (31.94)	13 (14.77)/5 (5.68)	1.514 (0.507-4.519)	0.455
Sex								
Females	41 (28.47)/42 (29.17)	29 (32.95)/18 (20.45)	0.061 (0.292-1.256)	0.176	53 (36.81)/30 (20.83)	34 (38.64)/13 (14.77)	1.480 (0.678-3.231)	0.323
Males	31 (21.53)/30 (20.83)	23 (26.14)/18 (20.45)	0.809 (0.365-1.791)	0.601	38 (26.39)/23 (15.97)	33 (37.5)/8 (9.09)	0.401 (0.158-1.015)	0.54
Smoking								
Never	50 (34.72)/52 (36.11)	40 (45.45)/27 (30.68)	0.649 (0.348-1.211)	0.173	67 (46.53)/35 (24.31)	54 (61.36)/13 (14.77)	**0.461 (0.222-0.957)**	**0.035**
Ever	22 (15.28)/20 (13.89)	12 (13.64)/9 (10.23)	0.825 (0.287-2.370)	0.721	35 (24.31)/7 (4.86)	16 (18.18)/5 (5.68)	1.563 (0.430-5.682)	0.496
BMI (kg/m^2)								
≤24	15 (10.42)/17 (11.81)	27 (30.68)/24 (27.27)	0.784 (0.323-1.902)	0.591	20 (13.89)/12 (8.33)	42 (47.73)/9 (10.23)	0.357 (0.129-0.986)	0.47
>24	56 (38.89)/56 (38.89)	25 (28.41)/12 (13.64)	0.480 (0.220-1.049)	0.063	71 (49.31)/41 (28.47)	25 (28.41)/12 (13.64)	1.203 (0.547-2.647)	0.646

The statistically significant are in the bold. [OR]odds ratio; [CI]confidence were interval. [1]Adjusted for age, sex, smoking status and BMI.

Table 5. Associations between variant SNHG18 rs788321 and rs2115243 genotypes and clinic pathologic characteristics of T2DM

Variables	(AT+ AA) and TT for rs788321		Allelic odds ratios and 95% confidence intervals for rs788321		(AC + CC) and AA for rs2115243		Allelic odds ratios and 95% confidence intervals for rs2115243	
	AT + AA, n	TT, n	Adjusted OR (95% CI)[1]	*P* value	AC + CC, n	AA, n	Adjusted OR (95% CI)[1]	*P* value
HbA1c (mmol, mean ± SD)								
~6	5	5	1		4	6	1	
6.1~10	58	56	1.036 (0.284-3.774)	0.55	72	42	2.571 (0.686-9.637)	1.61
>10	9	11	0.818 (0.179-3.744)	0.73	15	5	4.500 (0.890-22.734)	0.069
Duration (years, mean ± SD)								
~1	14	17	1		25	8	1	
1.1~5	32	28	1.388 (0.581-3.314)	0.994	29	29	**0.320 (0.124-0.826)**	**0.018**
5.1~10	18	17	1.286 (0.488-3.390)	0.632	26	9	0.924 (0.308-2.775)	0.889
>10	8	10	0.971 (0.302-3.124)	0.078	11	7	0.503 (0.146-1.734)	0.276
Treatment, n (%)								
Diet therapy (only)	4	7	1		5	6	1	
Exercise therapy (only)	3	2	2.625 (0.300-22.998)	0.383	2	3	0.800 (0.093-6.848)	0.839
Oral hypoglycemic (only)	22	12	3.208 (0.779-13.215)	0.107	20	14	1.714 (0.436-6.742)	0.44
Insulin (only)	11	5	3.850 (0761-19.468)	0.103	9	7	1.543 (0.329-7.226)	0.582
Combined therapy	49	29	2.975 (0.797-10.975)	0.105	39	39	1.200 (0.338-4.261)	0.778

The statistically significant are in the bold. [OR]odds ratio; [CI]confidence interval. [1]Alleles in the order of rs788321, rs2271223 and rs2115243. Only common haplotypes (frequency ≥5%) are listed.

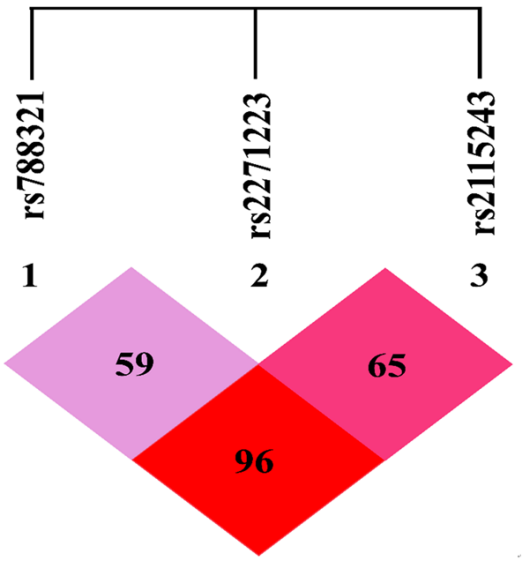

Figure 3. SHEsis showing linkage disequilibrium (LD) with r^2 values for three selected SNPs of SNHG18 gene. Note: dark red coloring displays strong LD, red displays intermediate LD and pink displays weak LD.

features of T2DM, including AbA1c and the treatment of T2DM (**Table 5**).

Association of SNHG18 haplotypes with risk and severity of T2DM

Linkage disequilibrium (LD) analysis showed weak LD between the three examined polymorphisms (D/=0.000-0.655, r^2=0.000-0.176) (**Figure 3**). Haplotype analyses were performed to further investigate combinational effects of these three polymorphisms on T2DM risk (**Table 6**). Using the most common A-A-A haplotype (alleles in the order of rs788321, rs2271223 and rs2115243) as a reference, A-G-C haplotype (*P*=0.007, adjusted OR=4.009, 95% CI=1.347-11.929), had a higher risk of T2DM. Haplotype analysis was performed to further evaluate the role of SNHG18 in T2DM susceptibility.

Discussion

Although lncRNA genetic polymorphisms were associated with susceptibility to many diseases, including esophageal squamous cell carcinoma (ESCC), gastric cancer, breast cancer, prostate cancer, lung cancer and papillary thyroid carcinoma [23-28]. And researches identified that lncRNAs were dys-expressed in T2DM [16, 17]. However, there were no related reports about the relevance between lncRNA genetic polymorphisms and T2DM. Our current study provides the first demonstration that association between lncRNA SNHG18 polymorphisms (rs788321, rs2271223 and rs2115243), and T2DM susceptibility in a southeast Chinese Han population. From our RNA sequencing data, we found lncRNA snhg8 which significant-

Table 6. Frequencies of estimated haplotypes between cases and controls and their risk prediction for T2DM

Haplotype[1]	Cases (%)	Controls (%)	Odds Ratio (95% CI)	*P* value
A-A-A	62.48 (21.7)	49.34 (28.0)	1	
A-G-A	28.31 (9.8)	10.38 (5.9)	1.734 (0.829-3.627)	0.139326
A-A-C	1.20 (0.4)	1.28 (0.7)	-	-
A-G-C	23.91 (8.3)	3.88 (2.2)	**4.009 (1.347-11.929)**	**0.007375**
T-A-A	14.6 (5.1)	6.78 (3.9)	1.328 (0.523-3.368)	0.549502
T-G-A	56.3 (19.6)	27.41 (15.6)	1.313 (0.795-2.167)	0.285890
T-A-C	101.19 (35.1)	76.93 (43.7)	0.693 (0472-1.018)	0.061123
T-G-C	0.02 (0.000)	0.00 (0.000)	-	-

The statistically significant are in the bold. [OR]odds ratio; [CI]confidence interval. [1]Alleles in the order of rs788321, rs2271223 and rs2115243. Only common haplotypes (frequency ≥5%) are listed.

ly down-regulated in T2DM db/db mice. Furthermore, we detected the expression of SNHG18 in human beings. Our data showed SNHG18 was down-expressed in T2DM patients. These results suggested SNHG18 may act as an important factor in T2DM.

In the present case-control study, we found an increased susceptibility of T2DM in subjects with mutant genotypes AT of rs788321 compared with carriers of the wild genotype TT (*P*=0.004, adjusted OR=2.602, 95% CI=1.359-4.979). Furthermore, a higher susceptibility of T2DM was also observed in individuals with the AT+AA genotype of rs788321 (*P*=0.030, adjusted OR=1.388, 95% CI=1.032-1.868). It suggested that SNHG18 were significantly associated with significantly increased T2DM susceptibility in the southwest Chinese Han population, modulating susceptibility to T2DM by influencing the expression level of SNHG18. These results indicated lncRNA polymorphisms may play a pathogenic role for T2DM.

Furthermore, a decreased susceptibility of T2DM was also observed in subjects with mutant genotypes AC in rs2115243 compared with carriers of the wild genotype AA (*P*=0.024, adjusted OR=0.465, 95% CI=0.239-0.905), indicating the AC genotypes might play a protective factor. And a higher susceptibility of T2DM was also observed in individuals with the AC+CC genotype (*P*=0.050, adjusted OR=0.729, 95% CI =0.531-1.000). In our study, we also found that individuals with the variant genotypes had a reduced risk of T2DM in non-smoker subjects (*P*=0.035, adjusted OR=0.461, 95% CI=0.222-0.957), compared with smoker subjects (*P*=0.496, adjusted OR=1.563, 95% CI=0.430-5.682). Our stratified analysis revealed that the protective effect was more pronounced in non-smokers, the observed adverse effects of AC genotypes may attribute to its genetic role in affecting smoking. In our stratified analysis to duration of T2DM in 1~5 years, those with the variant genotypes had significantly lower susceptibility of T2DM, while among other duration of T2DM, no statistical significance was noted. Our results indicated that the SNHG18 polymorphism rs2115243 (A>C) and no history of smoking may play a protective role for T2DM, and the shorter duration of T2DM may obstacle the development of T2DM patients.

In addition, our haplotype-based analysis results demonstrated that individuals with A-G-C haplotypes had a higher prevalence of T2DM when compared to those with the most common A-A-A haplotype, indicating that T allele of rs788321 and C allele of rs2115243 may also contribute to the risk of T2DM. Meanwhile, study proved that C-G-C-C, T-G-A-A, and T-A-A-A haplotypes of lncRNA H19 had a higher prevalence of coronary artery disease (CAD) [29]. These implied that the haplotypes of lncRNA associated with susceptibility of diseases and participate in the development of many diseases.

However, our study had many limitations. First of all, selection bias was existed because all subjects of disease groups and control groups were enrolled consecutively from hospitals during the same period and in the same area. But genotype distributions in the study were not deviate from the Hardy-Weinberg equilibrium in the controls and cases. Secondly, slight effects can't be detected because of the relatively small numbers and gene-environment interactions. Thirdly, because of the lack of clinical information, further analyses were prevented. In addition, demographic and personal information, such as smoking status and hypertension obtained by medical records and question-

naires may have a risk of false-positive or false-negative. Finally, the data of our study were unable to extrapolate for other regions and ethnic groups because which was performed in a southwest Chinese Han population.

In summary, we carried out the screening of T2DM-related lncRNA SNHG18 in T2DM and the association analysis between lncRNAs gene polymorphisms and T2DM patients for the first time. Our study showed that two potentially functional SNHG18 SNPs (rs788321 A>T and rs2115243 A>C) were significantly associated with T2DM in the southwest Chinese Han population. AT/AT+AA rs788321 genotypes were associated with significantly increased T2DM susceptibility and AC/AC+CC rs2115243 genotypes were associated with significantly decreased T2DM susceptibility. Furthermore, these two SNPs may have combined effects on the susceptibility of T2DM. Our study also demonstrated that the SNHG18 promoter rs788321 A>T polymorphism was functional, modulating susceptibility to T2DM by changing the expression levels of lncRNA. These results indicated SNHG18 polymorphisms may play a pathogenic role for T2DM. The rs788321 (A>T) and rs2115243 (A>C) polymorphism may change the translational efficiency, potentially lead to alter the structure of SNHG18, and affect the function of SNHG18. Nonetheless, the precise pathogenic mechanism of SNHG18 in T2DM still remains unclear, and further studies are required to support our presupposition.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (81270912) and the Scientific Research Foundation of Chongqing Medical University (201420).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Zheng Zhang, Department of Cell Biology and Medical Genetics, Chongqing Medical University, 1 Medical College Road, Chongqing 400016, China. Tel: +86 23 68485898; Fax: +86 23 68485555; E-mail: zhangzheng92@163.com

References

[1] Ali O. Genetics of type 2 diabetes. World J Diabetes 2013; 4: 114-123.

[2] Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract 2011; 94: 311-321.

[3] Boutayeb A, Boutayeb S. The burden of non-communicable diseases in developing countries. Int J Equity Health 2005; 4: 2.

[4] Booth GL, Kapral MK, Fung K, Tu JV. Relation between age and cardiovascular disease in men and women with diabetes compared with non-diabetic people: a population-based retrospective cohort study. Lancet 2006; 368: 29-36.

[5] Wolf G, Ritz E. Diabetic nephropathy in type 2 diabetes prevention and patient management. J Am Soc Nephrol 2003; 14: 1396-1405.

[6] Reddy MA, Chen Z, Park JT, Wang M, Lanting L, Zhang Q, Bhatt K, Leung A, Wu X, Putta S, Sætrom P, Devaraj S, Natarajan R. Regulation of Inflammatory Phenotype in Macrophages by a Diabetes-Induced Long Noncoding RNA. Diabetes 2014; 63: 4249-4261.

[7] Liu JY, Yao J, Li XM, Song YC, Wang XQ, Li YJ, Yan B, Jiang Q. Pathogenic role of lncRNA-MALAT1 in endothelial cell dysfunction in diabetes mellitus. Cell Death Dis 2014; 5: e1506.

[8] Spizzo R, Almeida MI, Colombatti A, Calin GA. Long non-coding RNAs and cancer: a new frontier of translational research? Oncogene 2012; 31: 4577-4587.

[9] Wapinski O, Chang, HY. Long noncoding RNAs and human disease. Trends Cell Biol 2011; 21: 354-361.

[10] Tsai MC, Spitale RC, Chang HY. Long intergenic noncoding RNAs: new links in cancer progression. Cancer Res 2011; 71: 3-7.

[11] Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010; 464: 1071-1076.

[12] Ponting CP, Oliver PL, Reik W. Evolution and functions of long non-coding RNAs. Cell 2009; 136: 629-641.

[13] Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol 2013; 20: 300-307.

[14] Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012; 22: 1775-1789.

[15] Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011; 43: 904-914.

[16] Morán I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, Arnes L, Nakić N, García-Hurtado J, Rodríguez-Seguí S, Pasquali L, Sauty-Colace C, Beucher A, Scharfmann R, van Arensbergen J, Johnson PR, Berry A, Lee C, Harkins T, Gmyr V, Pattou F, Kerr-Conte J, Piemonti L, Berney T, Hanley N, Gloyn AL, Sussel L, Langman L, Brayman KL, Sander M, McCarthy MI, Ravassard P, Ferrer J. Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab 2012; 16: 435-448.

[17] Carter G, Miladinovic B, Patel AA, Deland L, Mastorides S, Patel NA. Circulating long noncoding RNA GAS5 levels are correlated to prevalence of type 2 diabetes mellitus. BBA Clin 2015; 4: 102-107.

[18] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2013; 36 Suppl 1: S67-S74.

[19] Shi YY, He L. SHEsis, a powerful software platform for analyses of linkage dis-equilibrium, haplotype construction, and genetic association at Polymorphism loci. Cell Res 2005; 15: 97-98.

[20] Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, Diekhans M, Dreszer TR, Giardine BM, Harte RA, Hillman-Jackson J, Hsu F, Kirkup V, Kuhn RM, Learned K, Li CH, Meyer LR, Pohl A, Raney BJ, Rosenbloom KR, Smith KE, Haussler D, Kent WJ. The UCSC genome browser database: update 2011. Nucleic Acids Res 2011; 39: D876-D882.

[21] The International HapMap Consortium. The International HapMap Project. Nature 2003; 426: 789-796.

[22] Mathelier A, Forms O, Arenillas DJ, Chen CY, Denay G, Lee J, Shi W, Shyr C, Tan G, Worsley-Hunt R, Zhang AW, Parcy F, Lenhard B, Sandelin A, Wasserman WW. A major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res 2016; 44: D110-D115.

[23] Wu H, Zheng J, Deng J, Hu M, You Y, Li N, Li W, Lu J, Zhou Y. A genetic polymorphism in lincRNA-uc003opf.1 is associated with susceptibility to esophageal squamous cell carcinoma in Chinese populations. Carcinogenesis 2013; 34: 2908-2917.

[24] Yang C, Tang R, Ma X, Wang Y, Luo D, Xu Z, Zhu Y, Yang L. Tag SNPs in long non-coding RNA H19 contribute to susceptibility to gastric cancer in the Chinese Han population. Oncotarget 2015; 6: 15311-15320.

[25] Zhang Z, Zhu Z, Zhang B, Li W, Li X, Wu X, Wang L, Fu L, Fu L, Dong JT. Frequent mutation of rs13281615 and its association with PVT1 expression and cell proliferation in breast cancer. J Genet Genomics 2014; 41: 187-195.

[26] Gao W, Zhu M, Wang H, Zhao S, Zhao D, Yang Y, Wang ZM, Wang ZM, Wang F, Yang ZJ, Lu X, Wang LS. Association of polymorphisms in long non-coding RNA H19 with coronary artery disease risk in a Chinese population. Mutat Res 2015; 772: 15-22.

[27] Cao DL, Gu CY, Zhu Y, Dai B, Zhang HL, Shi GH, Shen YJ, Zhu YP, Ma CG, Xiao WJ, Qin XJ, Lin GW, Ye DW. Polymorphisms at long non-coding RNAs and prostate cancer risk in an eastern Chinese population. Prostate Cancer Prostatic Dis 2014; 17: 315-319.

[28] Jendrzejewski J, He H, Radomska HS, Li W, Tomsic J, Liyanarachchi S, Davuluri RV, Nagy R, de la Chapelle A. The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. Proc Natl Acad Sci U S A 2012; 109: 8646-8651.

[29] Gong WJ, Yin JY, Li XP, Fang C, Xiao D, Zhang W, Zhou HH, Li X, Liu ZQ. Association of well-characterized lung cancer lncRNA polymorphisms with lung cancer susceptibility and platinum-based chemotherapy response. Tumour Biol 2016; 37: 8349-8358.

[30] Liu O, Li J, Xin Y, Qin Y, Li H, Gong M, Liu Y, Wang X, Li J, Zhang H. Association of MMP-2 gene haplotypes with thoracic aortic dissection in Chinese Han population. BMC Cardiovasc Disord 2016; 16: 11.