ORIGINAL RESEARCH

Association Between Single Nucleotide Polymorphisms in *CDKAL1* and *HHEX* and Type 2 Diabetes in Chinese Population

Chuanyin Li[1],*
Keyu Shen[2],*
Man Yang[3]
Ying Yang[3]
Wenyu Tao[3]
Siqi He[4]
Li Shi[1]
Yufeng Yao [1]
Yiping Li[3]

[1]Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming, Yunnan 650118, People's Republic of China; [2]Department of Medicine, Dentistry and Healthy Science, The University of Melbourne, Melbourne, VIC 3010, Australia; [3]Department of Endocrinology and Metabolism, The Second People's Hospital of Yunnan Province, & the Affiliated Hospital of Yunnan University, Kunming, Yunnan 650021, People's Republic of China; [4]School of Clinical Medicine, Dali University, Dali, Yunnan 671000, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yufeng Yao
Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming, 650118 Yunnan, People's Republic of China
Email yufeng_yao@imbcams.com.cn

Yiping Li
Department of Endocrinology and Metabolism, The Second People's Hospital of Yunnan Province & the Affiliated Hospital of Yunnan University, Kunming, 650021 Yunnan, People's Republic of China
Email yyflyp@aliyun.com

Purpose: Type 2 diabetes mellitus (T2DM) has a high global prevalence, and the interaction of environmental factors and genetic factors may contribute to the risk of T2DM. We aimed to investigate the association between T2DM and the single nucleotide polymorphisms (SNPs) in genes (*CDKAL1* and *HHEX*) associated with insulin secretion.

Subjects and Methods: T2DM (n=1,169) and nondiabetic (NDM) (n=1,277) subjects were enrolled and the eight SNPs in *CDKAL1* and *HHEX* genes associated with insulin secretion were genotyped in a Chinese population using MassARRAY. Then, the association of these SNPs with T2DM was analyzed.

Results: Our results revealed that four SNPs (rs4712524, rs10946398, rs7754840 in *CDKAL1*, and rs5015480 in *HHEX*) showed significantly different distributions between the T2DM and NDM groups ($P<0.00625$). The G allele of rs4712524 ($P=0.004$, OR=1.184; 95% CI=1.057–1.327), C allele of rs10946398 ($P<0.001$, OR=1.247; 95% CI=1.112–1.398), and C allele of rs775480 in *CDKAL1* ($P<0.001$, OR=1.229; 95% CI=1.096–1.387) functioned as risk alleles of T2DM. The C allele of rs5015480 in *HHEX* ($P<0.001$, OR=1.295; 95% CI=1.124–1.493) was also the risk factor for T2DM. The haplotype analysis revealed that *CDKAL1* haplotype rs4712524G-rs10946398C-rs7754840C-rs9460546G ($P=0.001$, OR=1.210; 95% CI=1.076–1.360) and *HHEX* haplotype rs1111875C-rs5015480C ($P<0.001$, OR=1.364; 95% CI=1.180–1.576) were the risk factors of T2DM.

Conclusion: Our results revealed that genetic variations in *CDKAL1* and *HHEX* were associated with T2DM susceptibility in Chinese population.

Keywords: association study, single nucleotide polymorphisms, *CDKAL1*, *HHEX*, type 2 diabetes, Chinese population

Introduction

Type 2 diabetes mellitus (T2DM) is a complex disease which is associated with micro- and macrovascular complications. It can bring damage to several organs which is characterized by insulin resistance in peripheral tissues and dysregulated insulin secretion by pancreatic β-cell. The annual prevalence of T2DM in 2019 was 463 million adults globally, and is estimated to increase to over 700 million by 2045.[1] In China, the prevalence of diabetes mellitus showed a rapid increasing and reached 10.9% in 2013.[2] The Chinese population show a lower insulin secretion level compared with the European population, which implies that the insulin secretion function of pancreatic β-cell is a critical factor for the development of T2DM in the Chinese population.[3]

Cyclin-dependent kinase 5 regulatory subunit associated protein 1-like 1 (*CDKAL1*) gene encodes cyclin-dependent kinase 5 regulatory subunit-associated protein 1 (CDK5RAP1)-like 1, which is homologous to the CDK5RAP1, inhibitor of the CDK5 kinase through the inhibition of the CDK5 activator p35.[4,5] CDK5 is a serine/threonine protein kinase which is involved in the glucose-dependent regulation of the insulin secretion; therefore, it plays an important role in the pathophysiology of β-cell dysfunction and predisposition to T2DM.[6,7] Hematopoietically expressed homeobox (*HHEX*) gene encodes a transcription factor which is involved in Wnt signaling pathway, a fundamental pathway required for cell growth and development.[8] It plays a vital role in the ventral pancreatic specification.[9] Decreased somatostatin levels in HHEX-deficient islets causes disrupted paracrine inhibition of insulin release from β-cells.[10] According to the functional mechanism and downstream effects of these two genes, it is reasonable to conclude that *CDKAL1* and *HHEX* are well-accepted pathogenesis-related key genes for T2DM.

The associations of single nucleotide polymorphism (SNP) in both *CDKAL1* and *HHEX* with insulin secretion and/or T2DM had been examined in many studies.[11,12] It is noted that not only are there different association results in different ethnic groups, but also there are different association results in the same population.[13] For example, the associations of SNPs (rs1111875, rs5015480, and rs7923837) in *HHEX* with T2DM were observed in Chinese population from Shanghai, but not from Beijing.[13] Thus, it is necessary to carry out additional association studies to reveal the real roles of SNPs in *CDKAL1* and *HHEX* in susceptibility of T2DM. Here, we conducted a case-control study to investigate whether the SNPs located on CDKAL1 (rs4712524, rs10946398, rs7754840, rs9460546, and rs7756992) and HHEX (rs1111875, rs5015480, and rs7923837) were associated with the T2DM in the Chinese Han population.

Subjects and Methods
Ethics Statement
This study was carried out in accordance with the Helsinki Declaration of 1975, which was revised in 2008. The protocol employed by this investigation was approved by the Institutional Review Board of the Second People's Hospital of Yunnan Province before investigating. Each recruiter provided informed consent before participating in the study.

Subjects
A total of 1,169 T2DM patients and 1,277 healthy volunteers were enrolled in this study. All participants (T2DM and NDM) self-reported to be of Han ethnicity. The 1,169 patients of T2DM (742 males and 427 females) were diagnosed at the Second People's Hospital of Yunnan Province between January 2018 and December 2019 in accordance with the diagnostic standard of T2DM that World Health Organization (WHO) criteria published in 1999 and American Diabetes Association (ADA) guidelines in 2020.[14] For the T2DM group, the inclusion criteria are: 1) The T2DM subjects should be diagnosed as diabetes mellitus (DM) by fasting plasma glucose and/or 2 hours postprandial plasma glucose; 2) the T2DM subjects should be with fasting plasma glucose ≥7.0 mmol/L and/or 2 hours postprandial plasma glucose ≥11.1 mmol/L two times. The exclusion criteria are: 1) gGestational diabetes mellitus and other specific types of diabetes were excluded by requiring history illness; 2) the fasting and 2 hours postprandial plasma concentrations of insulin and C peptides were detected to exclude subjects with type 1 diabetes; and 3) subjects younger than 40 were excluded if islet cell autoantibodies and the glutamic acid decarboxylase autoantibodies are positive. For the NDM group, the inclusion criteria are: 1) tThe NDM subjects should be measured by fasting plasma glucose and glycosylated hemoglobin (HbA1C); and 2) the NDM subjects should be with fasting plasma glucose <6.10 mmol/L and HbA1C <5.7%. The exclusion criteria are: 1) sSubjects with a family history of diabetes mellitus were excluded; and 2) subjects with hypertension or coronary heart disease were also excluded.

Laboratory Measurements
Venous blood samples were collected in the morning after the subjects had fasted for 12 hours. Fasting plasma glucose (FPG) levels were measured using the glucose oxidase method. The levels of total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), triglycerides (TG), and low-density lipoprotein cholesterol (LDL-C) were measured by enzymatic methods. HbA1C was measured by immunoturbidimetry. All laboratory measurements were performed using a HITACHI 7600–020 Automatic Analyzer.

5114

Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy 2020:13

Selection and Genotyping of SNPs

We primarily focused on the two genes (*CDKAL1* and *HHEX*) which were shown to have a relationship with pathogenesis of T2DM. Five SNPs located in *CDKAL1* gene (rs4712524, rs10946398, rs7754840, rs9460546, and rs7756992) and three SNPs (rs1111875, rs5015480, and rs7923837) located in *HHEX* gene were selected.

Genomic DNA from peripheral blood was extracted using The QIAamp Blood Mini Kit (Qiagen, Hilden, Germany). MassARRAY Analyzer 4.0 (Agena, Inc) was used to genotype the eight SNPs from the two genes. The PCR primers were designed using the AssayDesigner 3.1 (Sequenom lnc., San Diego, CA, USA). Four microliters of the PCR master mix was mixed with 1 μL template DNA (25 ng/μL) in a 384-well plate. The PCR reaction setting was the same as those in our previous study.[15] Then 2 μL of shrimp alkaline phosphatase (SAP) was added into the PCR products per-well for removing the dNTP; the reaction conditions were set at: 37°C for 20 minutes and 85°C for 5 minutes. Next, 2 μL of the EXTEND Mix was added for single base extension using the following PCR cycle conditions: 1) 94°C for 30 seconds, 2) 94°C for 5 seconds, 3) 52°C for 5 seconds, 4) 80°C for 5 seconds, 5) 72–94°C for 3 minutes, and then steps 2–4 were repeated for 40 cycles with five repetitions of steps 3 and 4 per cycle. Resin purification was performed to 9 μL reaction products then the final products were transferred into a 384-well SpectroCHIP bioarray by MassARRAY Nanodispenser RS1000 machine (Agena, Inc, San Diego, CA, USA). The MALDI-TOF mass spectrometer (Agena, Inc) was used to translate the SpectroCHIP and the raw genotyping data was obtained using the TYPER 4.0 software.

Statistical Analysis

Student's *t*-tests were used to compare the ages, glucose, and lipid metabolic parameters (TC, HDL-C, TG, LDL-C, FPG, and HbA1C) between the T2DM and NDM groups while the gender distribution between the T2DM and control groups was tested using the Chi-square test. The Student's *t*-test and Chi-square test were performed using the SPSS 21. All polymorphic loci were tested for deviation from the Hardy-Weinberg equilibrium in the control group with a threshold of 0.05. Basic statistical analysis for the association between the alleles, genotypes, haplotypes, and T2DM were performed using the SHEsis software,[16,17] and the T2DM risk was estimated by the odds ratio (OR) with 95% confidence interval (95% CI). The distribution and differences in the haplotypes between the case and control groups were determined using the SHEsis software and the haplotypes with frequencies <0.03 were ignored during the analysis.[16,17] The association between T2DM and genotypes of these SNPs was analyzed using inheritance mode analysis by SNPStats software.[18] The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to determine the best fit model for each SNP.[18] The statistical power was calculated using PS Software.[19] The glucose and lipid metabolic parameters of the different genotypes in each SNP were compared with one-way analysis of variance. Multiple comparisons among the three genotypes were corrected by Tukey's method using GraphPad Prism software (Version 7.0). The significant threshold after Bonferroni correction for multiple comparisons was indicated by $P<0.00625(0.05/8)$ for each SNP.

Results

Subject Characteristics

The clinical characteristics and the glucose and lipid metabolic parameters of the enrolled subjects are presented in Table 1. The age or gender between the subjects from the T2DM and NDM groups showed no significant difference ($P>0.05$) as well as age between the male and female subjects from both groups. Otherwise, there were significant differences in the glucose and lipid metabolic parameters (TC, TG, HDL-C, LDL-C, FPG, and HbA1C levels) between subjects from the T2DM and NDM groups, that statistically showed as $P<0.001$ (Table 1). To be more specific, compared to the subjects from the NDM group, the T2DM subjects exhibited dyslipidemia (higher levels of TC, TG, and LDL-C and lower levels of HDL-C).

Association Between the SNPs and T2DM

The assessment of the genotype frequencies of the eight SNPs in the NDM group had been tested to be in Hardy-Weinberg equilibrium ($P>0.05$) (Table 2). The allele and genotype frequencies of the eight SNPs in both the T2DM and NDM groups are listed in Table 2. Four SNPs (rs4712524, rs10946398, rs7754840 in *CDKAL1*, and rs5015480 in *HHEX*) showed significantly different allele frequencies between the T2DM and NDM groups ($P<0.00625$) (Table 2). The G allele of rs4712524

Table 1 Clinical Characteristics and Glucose and Lipid Metabolic Parameters of the Subjects Enrolled in the Present Study

	Nondiabetic Subject	Type 2 Diabetes	P
N	1,277	1,169	
Age (years)	52.660±10.516	52.440±12.228	0.634
Sex (M/F)	781/496	742/427	0.243
Age (M; years)	51.636±10.521	51.305±12.583	0.578
Age (F; years)	54.272±10.314	54.412±11.332	0.845
Total cholesterol (mmol/L)	4.515±1.049	4.807±1.076	<0.001
Triglycerides (mmol/L)	1.761±1.352	2.455±2.113	<0.001
High-density lipoprotein-cholesterol (mmol/L)	1.280±0.380	1.094±0.296	<0.001
Low-density lipoprotein-cholesterol (mmol/L)	2.624±0.846	2.788±0.975	<0.001
Fasting plasma glucose (mmol/L)	5.006±0.526	7.888±2.465	<0.001
HbA1C (%)	5.139±0.345	8.987±2.823	<0.001

(P=0.004, OR=1.184; 95% CI=1.057–1.327), C allele of rs10946398 (P<0.001, OR=1.247; 95% CI=1.112–1.398), and C allele of rs7754840 (P<0.001, OR=1.229; 95% CI=1.096–1.378) in *CDKAL1*, and the C allele of rs5015480 of in *HHEX* (P<0.001, OR=1.295; 95% CI=1.124–1.493) functioned as the risk alleles of T2DM. The other four SNPs, rs9460546 and rs7756992 in *CDKAL1*, and rs1111875 and rs7923837 in *HHEX* did not show any significant association.

Association of the Haplotypes of the SNPs with T2DM

Linkage disequilibrium (LD) among the SNPs was also estimated, where the LD coefficient "D" was calculated using the SHEsis software. The *D'* values are defined in the range [−1, 1], with a value of "1" indicating perfect disequilibrium. A *D'* value above 0.8 indicated the existence of different loci in the LD. The LD of the five SNPs (rs4712524, rs10946398, rs7754840, rs9460546, and rs7756992) in *CDKAL1* was estimated. The *CDKAL1* rs4712524-rs10946398-rs7754840-rs9460546 haplotype was constructed with *D'*>0.9 (Supplementary Table 1). The haplotype analysis revealed that the frequency distributions of the haplotypes rs4712524-rs10946398-rs7754840-rs9460546 in the T2DM and NDM groups had statistically significant differences (P=0.001) (Table 3). This result indicated that the rs4712524G-rs10946398C-rs7754840C-rs9460546G (OR=1.210; 95% CI=1.076–1.360) was associated with a higher risk of

T2DM development, while rs4712524A-rs10946398A-rs7754840G-rs9460546T functioned as a protector against the development of T2DM (OR=0.827; 95% CI=0.735–0.930) (Table 3). The *HHEX* rs1111875-rs5015480 haplotype analysis with *D'*>0.9 showed that rs1111875C-rs5015480C was associated with development of T2DM (P<0.001, OR=1.364; 95% CI=1.180–1.576) (Table 4 and Supplementary Table 2).

Mode of Inheritance Analysis

The inheritance model (codominant, dominant, recessive, overdominant, and log-additive) of these SNPs was constructed using SNPStats. The best fit inheritance model for rs4712524, rs10946398, and rs7754840 in *CDKAL1* and rs5015480 in *HHEX* were all log-additive (with the lowest AIC and BIC). For rs4712524, 2G/G+G/A genotype was associated with higher risk of T2DM development (P=0.003, OR=1.19; 95% CI=1.06–1.33) (Table 5). For rs10946398, the 2C/C+C/A genotype was the risk factor compared to AA genotype (P<0.001, OR=1.26; 95% CI=1.12–1.42) (Table 6). For rs7754840, the 2C/C+C/G genotype was associated with higher risk of T2DM development (P<0.001, OR=1.24; 95% CI=1.11–1.40) (Table 7). For rs5015480 2C/C+C/G was the risk genotype (P<0.001, OR=1.30; 95% CI=1.13–1.51) (Table 8). The other SNPs did not exhibit any association with T2DM in the given population (Supplementary Table 3–6).

Table 2 Comparison of Genotypic and Allelic Distribution of SNPs (Rs4712524, Rs10946398, Rs7754840, Rs9460546, and Rs7756992) in CDKAL1 and SNPs (Rs111875, Rs5015480, and Rs7923837) in HHEX

SNPs	Allele		Chi²	P	Odds Ratio [95% CI]	Genotype			Chi²	P	H-W
rs4712524	A (freq)	G (freq)				A/A (freq)	A/G (freq)	G/G (freq)			
T2DM	1324 (0.566)	1,014 (0.434)	8.462	0.004	1.184 [1.057–1.327]	375 (0.321)	574 (0.491)	220 (0.188)	8.468	0.015	0.989
NDM	1551 (0.607)	1,003 (0.393)				470 (0.368)	611 (0.478)	196 (0.153)			0.911
rs10946398	C (freq)	A (freq)				C/C (freq)	C/A (freq)	A/A (freq)			
T2DM	996 (0.426)	1,342 (0.574)	14.233	<0.001	1.247 [1.112–1.398]	208 (0.178)	580 (0.496)	381 (0.326)	15.683	<0.001	0.619
NDM	953 (0.373)	1,601 (0.627)				162 (0.127)	629 (0.493)	486 (0.381)			0.058
rs7754840	C (freq)	G (freq)				C/C (freq)	C/G (freq)	G/G (freq)			
T2DM	999 (0.427)	1,339 (0.573)	12.421	<0.001	1.229 [1.096–1.378]	208 (0.178)	583 (0.499)	378 (0.323)	13.541	0.001	0.516
NDM	965 (0.378)	1,589 (0.622)				167 (0.131)	631 (0.494)	479 (0.375)			0.068
rs9460546	T (freq)	G (freq)				T/T (freq)	T/G (freq)	G/G (freq)			
T2DM	1,341 (0.574)	997 (0.426)	4.597	0.032	0.883 [0.788–0.989]	392 (0.335)	557 (0.476)	220 (0.188)	4.530	0.104	0.374
NDM	1,542 (0.604)	1,012 (0.396)				471 (0.369)	600 (0.470)	206 (0.161)			0.520
rs7756992	A (freq)	G (freq)				A/A (freq)	A/G (freq)	G/G (freq)			
T2DM	1,117 (0.478)	1,221 (0.522)	5.771	0.016	0.871 [0.779–0.975]	267 (0.228)	583 (0.499)	319 (0.273)	6.479	0.039	0.984
NDM	1,308 (0.512)	1,246 (0.488)				325 (0.255)	658 (0.515)	294 (0.230)			0.266
rs111875	T (freq)	C (freq)				T/T (freq)	T/C (freq)	C/C (freq)			
T2DM	1,640 (0.701)	698 (0.299)	4.703	0.030	0.871 [0.770–0.987]	577 (0.494)	486 (0.416)	106 (0.091)	6.340	0.042	0.800
NDM	1,863 (0.729)	691 (0.271)				669 (0.524)	525 (0.411)	83 (0.065)			0.137
rs5015480	C (freq)	T (freq)				C/C (freq)	C/T (freq)	T/T (freq)			
T2DM	504 (0.216)	1,834 (0.784)	12.815	<0.001	1.295 [1.124–1.493]	51 (0.044)	402 (0.344)	716 (0.612)	13.363	0.001	0.565
NDM	447 (0.175)	2,107 (0.825)				33 (0.026)	381 (0.298)	863 (0.676)			0.236
rs7923837	A (freq)	G (freq)				A/A (freq)	A/G (freq)	G/G (freq)			
T2DM	1,755 (0.751)	583 (0.249)	1.794	0.181	0.914 [0.802–1.042]	656 (0.561)	443 (0.379)	70 (0.060)	2.822	0.244	0.674
NDM	1,959 (0.767)	595 (0.233)				740 (0.579)	479 (0.375)	58 (0.045)			0.077

Table 3 The CDKAL1 Rs4712524-Rs10946398-Rs7754840-Rs9460546 Haplotype Analysis Between the NDM and T2DM Group

Haplotypes	T2DM(Freq)	NDM(Freq)	Chi^2	P	Odds Ratio [95% CI]
AAGT	1,287.58 (0.551)	1,494.20 (0.585)	10.099	0.001	0.827 [0.735–0.930]
GCCG	954.53 (0.408)	915.83 (0.359)	10.099	0.001	1.210 [1.076–1.360]

Note: All those frequencies <0.03 will be ignored in analysis.

Table 4 The HHEX Rs1111875-Rs5015480 Haplotype Analysis Between the NDM and T2DM Group

Haplotypes	T2DM(Freq)	NDM(Freq)	Chi^2	P	Odds Ratio [95% CI]
TT	1,629.00 (0.697)	1,831.38 (0.717)	4.006	0.045	0.881 [0.777–0.997]
CC	493.00 (0.211)	415.39 (0.163)	17.700	<0.001	1.364 [1.180–1.576]
CT	205.00 (0.088)	275.61 (0.108)	6.077	0.014	0.788 [0.651–0.953]

Note: All those frequencies <0.03 will be ignored in analysis.

Table 5 Different Inheritance Models Analysis of the Rs4712524 in CDKAL1 Between the NDM and T2DM Group

Model	Genotype	NDM	T2DM	OR (95% CI)	P	AIC	BIC
Codominant	A/A	470 (36.8%)	375 (32.1%)	1	0.013	3,385.9	3,414.9
	G/A	611 (47.9%)	574 (49.1%)	1.18 (0.99–1.42)			
	G/G	196 (15.3%)	220 (18.8%)	1.41 (1.12–1.79)			
Dominant	A/A	470 (36.8%)	375 (32.1%)	1	0.012	3,386.3	3,409.5
	G/A-G/G	807 (63.2%)	794 (67.9%)	1.24 (1.05–1.47)			
Recessive	A/A-G/A	1,081 (84.7%)	949 (81.2%)	1	0.023	3,387.4	3,410.6
	G/G	196 (15.3%)	220 (18.8%)	1.28 (1.03–1.58)			
Overdominant	A/A-G/G	666 (52.1%)	595 (50.9%)	1	0.500	3,392.1	3,415.3
	G/A	611 (47.9%)	574 (49.1%)	1.06 (0.90–1.24)			
Log-additive	—	—	—	1.19 (1.06–1.33)	0.003	3,383.9	3,407.1

Table 6 Different Inheritance Models Analysis of the Rs10946398 in CDKAL1 Between the NDM and T2DM Group

Model	Genotype	NDM	T2DM	OR (95% CI)	P	AIC	BIC
Codominant	A/A	486 (38.1%)	381 (32.6%)	1	$3.0×10^{-4}$	3,378.6	3,407.6
	C/A	629 (49.3%)	580 (49.6%)	1.18 (0.99–1.41)			
	C/C	162 (12.7%)	208 (17.8%)	1.65 (1.29–2.10)			
Dominant	A/A	486 (38.1%)	381 (32.6%)	1	0.0038	3,384.2	3,407.4
	C/A-C/C	791 (61.9%)	788 (67.4%)	1.28 (1.08–1.51)			
Recessive	A/A-C/A	1,115 (87.3%)	961 (82.2%)	1	$4.0×10^{-4}$	3,380.2	3,403.4
	C/C	162 (12.7%)	208 (17.8%)	1.49 (1.19–1.86)			
Overdominant	A/A-C/C	648 (50.7%)	589 (50.4%)	1	0.810	3,392.5	3,415.8
	C/A	629 (49.3%)	580 (49.6%)	1.02 (0.87–1.20)			
Log-additive	—	—	—	1.26 (1.12–1.42)	<0.001	3,377.5	3,400.7

Table 7 Different Inheritance Models Analysis of the Rs7754840 in CDKAL1 Between the NDM and T2DM Group

Model	Genotype	NDM	T2DM	OR (95% CI)	P	AIC	BIC
Codominant	G/G G/C C/C	479 (37.5%) 631 (49.4%) 167 (13.1%)	378 (32.3%) 583 (49.9%) 208 (17.8%)	1 1.18 (0.99–1.41) 1.59 (1.24–2.03)	0.001	3,380.7	3,409.7
Dominant	G/G G/C-C/C	479 (37.5%) 798 (62.5%)	378 (32.3%) 791 (67.7%)	1 1.27 (1.07–1.50)	0.0058	3,385.0	3,408.2
Recessive	G/G-G/C C/C	1,110 (86.9%) 167 (13.1%)	961 (82.2%) 208 (17.8%)	1 1.44 (1.15–1.80)	0.0012	3,382.1	3,405.3
Overdominant	G/G-C/C G/C	646 (50.6%) 631 (49.4%)	586 (50.1%) 583 (49.9%)	1 1.02 (0.87–1.20)	0.770	3,392.5	3,415.7
Log-additive	—	—	—	1.24 (1.11–1.40)	<0.001	3,379.3	3,402.5

Table 8 Different Inheritance Models Analysis of the Rs5015480 in HHEX Between the NDM and T2DM Group

Model	Genotype	NDM	T2DM	OR (95% CI)	P	AIC	BIC
Codominant	T/T C/T C/C	863 (67.6%) 381 (29.8%) 33 (2.6%)	716 (61.2%) 402 (34.4%) 51 (4.4%)	1 1.27 (1.07–1.51) 1.87 (1.19–2.93)	0.001	3,381.2	3,410.2
Dominant	T/T C/T-C/C	863 (67.6%) 414 (32.4%)	716 (61.2%) 453 (38.8%)	1 1.32 (1.12–1.56)	0.001	3,381.9	3,405.1
Recessive	T/T-C/T C/C	1,244 (97.4%) 33 (2.6%)	1,118 (95.6%) 51 (4.4%)	1 1.73 (1.10–2.69)	0.015	3,386.7	3,409.9
Overdominant	T/T-C/C C/T	896 (70.2%) 381 (29.8%)	767 (65.6%) 402 (34.4%)	1 1.23 (1.04–1.46)	0.016	3,386.8	3,410.0
Log-additive	—	—	—	1.30 (1.13–1.51)	<0.001	3,379.5	3,402.7

Association of Genotypes of the SNPs with Metabolic Phenotypes

In the NDM group, no significant associations of these SNPs with glucose and lipid metabolic parameters, including FPG, TC, HDL-C, TG, LDL-C, and HbA1C, were observed (Supplementary Table 7).

Discussion

CDKAL1 and HHEX are well-accepted pathogenesis-related key genes for T2DM, they play a pivotal role in regulating the insulin secretion function of pancreatic β-cells. This study evaluated eight SNPs on *CDKAL1* and *HHEX* genes to determine the significant association with T2DM. Of those, four SNPs (rs4712524, rs10946398, and rs7754840 in *CDKAL1*, rs5015480 in *HHEX*) were associated with T2DM in the Chinese population.

CDKAL1 expression in human pancreatic β-cells participates in the correlation of CDKAL1- and CDK5-mediated pathways.[20] It has been shown that CDKAL1 could increase insulin secretion in pancreatic β-cells by inhibiting CDK5.[4,6,7] Recently, many studies have investigated the association between SNPs (rs7754840, rs4712524, rs10946398, rs7756992, and rs10811661) in the *CDKAL1* gene and T2DM in different populations around the world.[21–33] However, the association results of some SNPs showed a difference in different populations,[23,34] even in the same population.[26,35] In the current study, we found that rs7754840 was associated with T2DM in a Chinese population. Our results were consistent with other population results,[27,–29–33,36] especially for the Chinese population (Uyghur and Han).[37–39] In 2014, Klimentidis et al[40] reported that rs7754840 was significantly associated with HbA1c in a Yup'ik

population (P<0.001). And the rs7754840 also has been reported to be associated with decreased β-cell glucose sensitivity and insulin secretion.[11,41] These results indicated that rs7754840 could be associated with the T2DM through influencing the β-cell function and insulin secretion. However, in 2017, Nikitin et al[34] found the rs7754840 was not associated with T2DM in Russian. Lee et al[26] found rs7754840 was associated with risk of T2DM in Koreans in 2008. However, in 2012, Park et al[35] did not find the association between this SNP and T2DM in Koreans. These association results indicated that the different genetic background or other factors could influence the association results in different populations, even in the same population. This phenomenon was also found in the association of rs7756992 in a Chinese population. In 2010, Xu et al[42] reported an association of this SNP with T2DM in a Chinese population, which is similar with the results in other populations.[25–28,33] However, Dou et al[43] found there is no association between this SNP and T2DM in another Chinese population, which is consistent with the results in other populations.[28,39,44] In the current study, we also did not find an association between this SNP and T2DM in the Chinese population. The inconsistency was also found in the association of rs10946398 with T2DM, which was associated with T2DM in the current Chinese population, Chinese She,[45] and Russian[34] population, but not associated with T2DM in a Korean population.[25] In summary of the previous studies' results and our study, we found that these results did not show a clear conclusion of the association between these SNPs and T2DM. In 2015, Locke et al[46] reported that SNPs in *CDKAL1* might show an effect on the expression of CDKAL1. Thus, functional studies are needed to confirm the association and reveal the value of the association in clinical application.

HHEX gene encodes a transcription factor which is involved in a fundamental pathway, the Wnt signaling pathway.[10] In 2014, McKenna et al[47] found that misregulated HHEX expression within the diabetic islet might contribute to disrupted paracrine control of insulin secretion in T2DM, leading to accelerated β-cell exhaustion and failure. In the current study, we found that rs5015480C in *HHEX* gene might be associated with high risk of T2DM in the Chinese population. Our result is consistent with the results in German,[23] Korean,[25] and other Chinese populations,[13,38,48] including Chinese She population (the frequencies of rs5015480C were T2DM 0.437 vs NGT 0.403).[45] The rs5015480, which is located upstream of the *HHEX* gene has been found to be

associated with decreased insulin response and β-cell glucose sensitivity.[11] In 2014, Klimentidis et al[40] found that rs5015480 was significantly associated with a combined fasting glucose and HbA1c measure (P<0.001) and with HOMA-B (P<0.001) among the Yup'ik population. These results indicated that rs5015480 could be associated with the T2DM through influencing the β-cell glucose sensitivity.[11,41] Several studies showed that rs1111875 in *HHEX* gene was related to the risk of T2DM in European[11,49,50] and Chinese[13,48] populations. In the current study, our results showed that rs1111875T was the protective factor of T2DM before Bonferroni correction, which is consistent with results of previous studies.[11,–13,–48–50] However, there was no significant difference between T2DM and NDM after Bonferroni correction. If the sample size will be increased, the rs1111875 could be associated with the T2DM in the current Chinese population.

The haplotype is more powerful than single SNP in detecting the association between SNPs and complex diseases. In the current study, we observed that the rs4712524G-rs10946398C-rs7754840C-rs9460546G haplotype in *CDKAL1* gene was associated with a higher risk of T2DM (P=0.001, OR=1.210; 95% CI=1.076–1.360), which is consistent with the single SNP result. Moreover, the *HHEX* haplotype rs1111875C–rs5015480C was associated with development of T2DM which is also consistent with the single SNP result. Therefore, the association between these haplotypes and T2DM could provide powerful evidence in assessing the underlying genetic preposition in T2DM.

In the current study, our data just provided the association data between these SNPs and T2DM, however, the mechanisms still need to be verified through functional study, which is one of the limitations of the current study. In addition, some clinical information, such as weight, Body mass index (BMI), the treatment used in patients with T2DM, family history of T2DM, and metabolic-carbohydrate control was important. However, we did not collect these kinds of information in the current study. Thus, the lack of these kinds of information was the other limitation of the current study.

Conclusion

In the current study, we revealed four SNPs (rs4712524, rs10946398, and rs7754840 in *CDKAL1*, rs5085140 in *HHEX*) and the haplotype in *HHEX* and *CDKAL1* were associated with T2DM in a Chinese population. The statistical power of the four SNPs was 0.830, 0.965, 0.942, and 0.946, respectively, which indicates that our results are

reliable. Our results from an association study would provide reference data for revealing the pathogenesis of T2DM and finding the new biomarkers of T2DM, which could be utilized for the prevention, diagnosis, and treatment of T2DM.

Abbreviations

T2DM, type 2 diabetes mellitus; SNP, single nucleotide polymorphisms; NDM, nondiabetic; CDKAL1, cyclin-dependent kinase 5 regulatory subunit associated protein 1-like 1; HHEX, hematopoietically expressed homeobox; FPG, fasting plasma glucose; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides; LDL-C, low-density lipoprotein cholesterol; OR, odds ratio; LD, linkage disequilibrium; AIC, aAkaike information criterion; BIC, Bayesian information criterion.

Data Sharing Statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

The study was approved by The Institutional Review Board of the Second People's Hospital of Yunnan Province before the commencement of the investigation. Moreover, the principles stated in the Helsinki Declaration of 1975 revised in 2008 were followed by the protocol adopted in this investigation. Each participant signed an informed consent form.

Funding

Disclosure

The authors have declared that no competing interests exist.

References

1. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157:107843. doi:10.1016/j.diabres.2019.107843
2. Wang L, Gao P, Zhang M, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. *JAMA*. 2017;317 (24):2515. doi:10.1001/jama.2017.7596
3. Hu C, Jia W. Diabetes in China: epidemiology and genetic risk factors and their clinical utility in personalized medication. *Diabetes*. 2018;67(1):3–11. doi:10.2337/dbi17-0013
4. Ching Y-P, Pang ASH, Lam W-H, Qi RZ, Wang JH. Identification of a neuronal Cdk5 activator-binding protein as Cdk5 inhibitor. *J Biol Chem*. 2002;277(18):15237–15240. doi:10.1074/jbc.C200032200
5. Liu KC, Leuckx G, Sakano D, et al. Inhibition of Cdk5 promotes beta-cell differentiation from ductal progenitors. *Diabetes*. 2018;67:58–70. doi:10.2337/db16-1587
6. Wei F-Y, Nagashima K, Ohshima T, et al. Cdk5-dependent regulation of glucose-stimulated insulin secretion. *Nat Med*. 2005;11 (10):1104–1108. doi:10.1038/nm1299
7. Ubeda M, Rukstalis JM, Habener JF. Inhibition of cyclin-dependent kinase 5 activity protects pancreatic beta cells from glucotoxicity. *J Biol Chem*. 2006;281(39):28858–28864. doi:10.1074/jbc.M604690200
8. Kommoju UJ, Samy SK, Maruda J, Irgam K, Reddy BM. Association of CDKAL1, CDKN2A/B & HHEX gene polymorphisms with type 2 diabetes mellitus in the population of Hyderabad, India. *Indian J Med Res*. 2016;143:455. doi:10.4103/0971-5916.184303
9. Bort R, Martinez-Barbera JP, Beddington RS, Zaret KS. Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development*. 2004;131:797–806. doi:10.1242/dev.00965
10. Foley AC, Mercola M. Heart induction by Wnt antagonists depends on the homeodomain transcription factor Hex. *Genes Dev*. 2005;19:387–396. doi:10.1101/gad.1279405
11. Pascoe L, Tura A, Patel SK, et al. Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. *Diabetes*. 2007;56:3101–3104. doi:10.2337/db07-0634
12. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet*. 2007;39:770–775. doi:10.1038/ng2043
13. Wu Y, Li H, Loos RJF, et al. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes*. 2008;57(10):2834–2842. doi:10.2337/db08-0047
14. American Diabetes, A. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes-2020. *Diabetes Care*. 2020;43: S14–S31. doi:10.2337/dc20-S002
15. Li Y, et al. Association of single nucleotide polymorphisms of miRNAs involved in the GLUT4 pathway in T2DM in a Chinese population. 2019;7:e907. doi:10.1002/mgg3.907
16. Li Z, Zhang Z, He Z, et al. A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis. *Cell Res*. 2009;19(4):519–523. doi:10.1038/cr.2009.33
17. Yong YY, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res*. 2005;15(2):97–98. doi:10.1038/sj.cr.7290272
18. Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics*. 2006;22 (15):1928–1929. doi:10.1093/bioinformatics/btl268

19. Dupont WD, Plummer WD Jr. Power and sample size calculations for studies involving linear regression. *Control Clin Trials*. 1998;19 (6):589–601. doi:10.1016/S0197-2456(98)00037-3

20. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. 2007;316(5829):1336–1341. doi:10.1126/science.1142364

21. Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science (New York, N Y)*. 2007;316(5829):1331–1336. doi:10.1126/science.1142358

22. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med*. 2008;359(21):2220–2232. doi:10.1056/NEJMoa0801869

23. Herder C, Rathmann W, Strassburger K, et al. Variants of the PPARG, IGF2BP2, CDKAL1, HHEX, and TCF7L2 Genes Confer Risk of Type 2 Diabetes Independently of BMI in the German KORA Studies. *Hormone and Metabolic Research = Hormon- Und Stoffwechselforschung = Hormones Et Metabolisme*. 2008;40 (10):722–726. doi:10.1055/s-2008-1078730

24. Loganadan NK, Huri HZ, Vethakkan SR, Hussein Z. Genetic markers predicting sulphonylurea treatment outcomes in type 2 diabetes patients: current evidence and challenges for clinical implementation. *Pharmacogenomics J*. 2016;16(3):209–219. doi:10.1038/tpj.2015.95

25. Ryoo H, Woo J, Kim Y, Lee C. Heterogeneity of genetic associations of CDKAL1 and HHEX with susceptibility of type 2 diabetes mellitus by gender. *Eur j Human Genetics*. 2011;19(6):672–675. doi:10.1038/ejhg.2011.6

26. Lee Y-H, Kang ES, Kim SH, et al. Association between polymorphisms in SLC30A8, HHEX, CDKN2A/B, IGF2BP2, FTO, WFS1, CDKAL1, KCNQ1 and type 2 diabetes in the Korean population. *J Hum Genet*. 2008;53(11–12):991–998. doi:10.1007/s10038-008-0341-8

27. Tabara Y, Osawa H, Kawamoto R, et al. Genotype risk score of common susceptible variants for prediction of type 2 diabetes mellitus in Japanese: the Shimanami Health Promoting Program (J-SHIPP study). Development of type 2 diabetes mellitus and genotype risk score. *Metabolism*. 2011;60(11):1634–1640. doi:10.1016/j.metabol.2011.03.014

28. Tan JT, Ng DPK, Nurbaya S, et al. Polymorphisms identified through genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore. *J Clin Endocrinol Metab*. 2010;95(1):390–397. doi:10.1210/jc.2009-0688

29. Kommoju UJ, Samy S, Maruda J, et al. Association of CDKAL1, CDKN2A/B & HHEX gene polymorphisms with type 2 diabetes mellitus in the population of Hyderabad, India. *Indian J Med Res*. 2016;143(4):455–463. doi:10.4103/0971-5916.184303

30. Soltani G, Hatefi Z, Salehi A, et al. Pharmacogenomics of Sulfonylureas response in relation to rs7754840 Polymorphisms in Cyclin-Dependent Kinase 5 regulatory subunit-associated protein 1-like (CDKAL1) Gene in Iranian Type 2 diabetes patients. *Adv Biomed Res*. 2018;7(1):96. doi:10.4103/abr.abr_144_17

31. Vatankhah Yazdi K, Kalantar SM, Houshmand M, et al. SLC30A8, CDKAL1, TCF7L2, KCNQ1 and IGF2BP2 are associated with Type 2 diabetes mellitus in Iranian patients. *Diabetes Metabol Syndr Obesity*. 2020;13:897–906. doi:10.2147/dmso.s225968

32. Mansoori Y, Daraei A, Naghizadeh MM, Salehi R. Significance of a common variant in the CDKAL1 gene with susceptibility to type 2 diabetes mellitus in Iranian population. *Adv Biomed Res*. 2015;4:45. doi:10.4103/2277-9175.151256

33. Nemr R, Almawi AW, Echtay A, et al. Replication study of common variants in CDKAL1 and CDKN2A/2B genes associated with type 2 diabetes in Lebanese Arab population. *Diabetes Res Clin Pract*. 2012;95:e37–40. doi:10.1016/j.diabres.2011.11.002

34. Nikitin AG, Potapov VY, Brovkina OI, et al. Association of polymorphic markers of genes FTO, KCNJ11, CDKAL1, SLC30A8, and CDKN2B with type 2 diabetes mellitus in the Russian population. *PeerJ*. 2017;5:e3414. doi:10.7717/peerj.3414

35. Park SE, et al. Impact of common type 2 diabetes risk gene variants on future type 2 diabetes in the non-diabetic population in Korea. *J Hum Genet*. 2012;57:265–268. doi:10.1038/jhg.2012.16

36. Kanthimathi S, Chidambaram M, Liju S, et al. Identification of genetic variants of gestational diabetes in South Indians. *Diabetes Technol Ther*. 2015;17:462–467. doi:10.1089/dia.2014.0349

37. Song M, Zhao F, Ran L, et al. The Uyghur population and genetic susceptibility to Type 2 diabetes: potential role for variants in CDKAL1, JAZF1, and IGF1 genes. *Omics*. 2015;19(4):230–237. doi:10.1089/omi.2014.0162

38. Han X, Luo Y, Ren Q, et al. Implication of genetic variants near SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, FTO, TCF2, KCNQ1, and WFS1 in type 2 diabetes in a Chinese population. *BMC Med Genet*. 2010;11(1):81. doi:10.1186/1471-2350-11-81

39. Bao XY, Peng B, Yang MS. Replication study of novel risk variants in six genes with type 2 diabetes and related quantitative traits in the Han Chinese lean individuals. *Mol Biol Rep*. 2012;39(3):2447–2454. doi:10.1007/s11033-011-0995-8

40. Klimentidis YC, Lemas DJ, Wiener HH, et al. CDKAL1 and HHEX are associated with type 2 diabetes-related traits among Yup'ik people. *J Diabetes*. 2014;6:251–259. doi:10.1111/1753-0407.12093

41. Chistiakov DA, Potapov VA, Smetanina SA, et al. The carriage of risk variants of CDKAL1 impairs beta-cell function in both diabetic and non-diabetic patients and reduces response to non-sulfonylurea and sulfonylurea agonists of the pancreatic KATP channel. *Acta Diabetol*. 2011;48:227–235. doi:10.1007/s00592-011-0299-4

42. Xu M, Bi Y, Xu Y, et al. Combined effects of 19 common variations on type 2 diabetes in Chinese: results from two community-based studies. *PLoS One*. 2010;5:e14022. doi:10.1371/journal.pone.0014022

43. Dou H-Y, Wang -Y-Y, Yang N, et al. Association between genetic variants and characteristic symptoms of type 2 diabetes: a matched case-control study. *Chin J Integr Med*. 2017;23(6):415–424. doi:10.1007/s11655-015-2290-3

44. El-Lebedy D, Ashmawy I. Common variants in TCF7L2 and CDKAL1 genes and risk of type 2 diabetes mellitus in Egyptians. *J Genetic Eng Biotechnol*. 2016;14:247–251. doi:10.1016/j.jgeb.2016.10.004

45. Chen G, Xu Y, Lin Y, et al. Association study of genetic variants of 17 diabetes-related genes/loci and cardiovascular risk and diabetic nephropathy in the Chinese She population. *J Diabetes*. 2013;5:136–145. doi:10.1111/1753-0407.12025

46. Locke JM, Wei FY, Tomizawa K, Weedon MN, Harries LW. A cautionary tale: the non-causal association between type 2 diabetes risk SNP, rs7756992, and levels of non-coding RNA, CDKAL1-v1. *Diabetologia*. 2015;58:745–748. doi:10.1007/s00125-015-3508-9

47. McKenna, et al. The diabetes gene Hhex maintains δ-cell differentiation and islet function. *Genes Develop*. 2014.

48. Zhou D-Z, Liu Y, Zhang D, et al. Variations in/nearby genes coding for JAZF1, TSPAN8/LGR5 and HHEX-IDE and risk of type 2 diabetes in Han Chinese. *J Hum Genet*. 2010;55(12):810–815. doi:10.1038/jhg.2010.117

49. Saxena R, Voight BF, Lyssenko V, et al.; Diabetes Genetics Initiative of Broad Institute of, H. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. 2007;316 (5829):1331–1336. doi:10.1126/science.1142358

50. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007;445 (7130):881–885. doi:10.1038/nature05616

Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy 2020:13   submit your manuscript | www.dovepress.com   **5123**

DovePress