Diabetologia (2008) 51:2134–2137
DOI 10.1007/s00125-008-1141-6

RESEARCH LETTER

# Positive association between variations in *CDKAL1* and type 2 diabetes in Han Chinese individuals

**Y. Liu · L. Yu · D. Zhang · Z. Chen · D. Z. Zhou · T. Zhao · S. Li · T. Wang · X. Hu · G. Y. Feng · Z. F. Zhang · L. He · H. Xu**

Received: 18 July 2008 / Accepted: 24 July 2008 / Published online: 3 September 2008


**Keywords** Association study · Genetic variations · Type 2 diabetes

**Abbreviations**
IFG impaired fasting glucose
IGT impaired glucose tolerance
SNP single nucleotide polymorphism

Y. Liu · L. Yu · D. Zhang · Z. Chen · D. Z. Zhou · T. Zhao · S. Li · T. Wang · L. He (✉) · H. Xu (✉)
Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences,
294 Taiyuan Road,
Shanghai 200031, People's Republic of China
e-mail: helin@bio-x.cn
e-mail: hexuc@yahoo.com.cn

Y. Liu · L. Yu · D. Zhang · Z. Chen · D. Z. Zhou · T. Zhao · S. Li · G. Y. Feng · L. He · H. Xu
Bio-X Center, Shanghai Jiao Tong University,
Shanghai, People's Republic of China

X. Hu
Department of Stomatology, Ruijin Hospital,
Luwan Branch,
Shanghai, People's Republic of China

Z. F. Zhang
Department of Epidemiology, UCLA School of Public Health,
Los Angeles, CA, USA

L. He
The Obstetrics and Gynecology Hospital, Fudan University,
Shanghai 200011, 419 Fangxie Road,
China

L. He
Institutes of Biomedical Sciences, Fudan University,
Shanghai, People's Republic of China

*To the Editor*: Type 2 diabetes is a complex disease that has become a rapidly growing public health problem in China. The molecular mechanisms and underlying genetic architecture of type 2 diabetes still need to be clarified. Genome-wide association studies have recently described novel type 2 diabetes susceptibility loci, including several previously unknown genomic regions, such as *CDKAL1* and *IGF2BP2* [1–4]. Following these discoveries, a replication study has confirmed a significant association between type 2 diabetes and these two genes in the Japanese population [5].

To investigate whether variations in *CDKAL1* and *IGF2BP2* also play a major role in type 2 diabetes susceptibility in China, we genotyped five representative variants (rs7754840, rs10946398 and rs7756992 in *CDKAL1*; rs4402960 and rs1470579 in *IGF2BP2*) in 4,189 Han Chinese participants, including 1,912 unrelated type 2 diabetic patients (785 men, 1,127 women; age 63.8±9 years), 236 individuals with impaired fasting glucose or impaired glucose tolerance (IFG/IGT; 83 men, 153 women; age 64±9 years) and 2,041 control individuals (635 men, 1,406 women; age 58±9 years). Diabetic and IFG/IGT participants were defined in accordance with WHO criteria. Controls with a fasting plasma glucose concentration of <6.1 mmol/l were enrolled from the same region. The clinical characteristics of the participants are summarised in Table 1. The study protocol was reviewed and approved by the Ethics Committee of the Shanghai Institute for Biological Sciences and all participants gave informed consent. High-molecular-weight genomic DNA was prepared from venous blood using the QuickGene 610L Automatic DNA/RNA Extraction System (Fijifilm, Tokyo, Japan). All five single nucleotide polymorphisms (SNPs) were genotyped using TaqMan technology on an ABI7900 system (Applied Biosystems, Foster City, CA, USA). Genotype data were

**Table 1** Clinical characteristics of the participants

| Characteristics | Type 2 diabetes | IFG/IGT | Control |
|---|---|---|---|
| Age (years) | 63.8±9 | 64±9 | 58.1±9 |
| BMI (kg/m$^2$) | 25.3±3.4 | 24.9±3.2 | 24.5±3.2 |
| WHR | 0.90±0.06 | 0.88±0.06 | 0.86±0.06 |
| Fasting blood glucose (mmol/l) | 8.1±2.9 | 5.6±0.68 | 4.8±0.64 |
| Triacylglycerol (mmol/l) | 1.9±1.6 | 1.9±1.2 | 1.6±1.1 |
| Cholesterol (mmol/l) | 4.5±0.9 | 4.5±0.8 | 4.5±0.9 |
| HDL-cholesterol (mmol/l) | 1.2±0.3 | 1.1±0.3 | 1.3±0.3 |
| LDL-cholesterol (mmol/l) | 2.8±0.8 | 2.8±0.7 | 2.8±0.7 |
| Systolic BP (mmHg) | 140.1±20.4 | 141.3±18.9 | 132±19.3 |
| Diastolic BP (mmHg) | 82.0±10.2 | 82.1±10.6 | 81.1±10.1 |
| $HbA_{1c}$ (%) | 7.5±1.6 | 6.2±0.5 | 5.8±0.6 |

Data are means±SD

obtained for more than 90% of the DNA samples and replicate quality control samples (32 samples) were genotyped with 100% concordance. The association of SNPs with type 2 diabetes was assessed by logistic regression after adjusting for sex, age and $\log_e$-transformed BMI ($\log_e$-BMI). For regression modelling in the additive model, individuals homozygous for the risk allele (1/1), heterozygous (1/0) and homozygous for the non-risk allele (0/0) were assigned a categorical variable according to genotype (2, 1 and 0, respectively). The dominant model was defined as 1/1 + 1/0 vs 0/0, and the recessive model as 1/1 vs 1/0 + 0/0. The statistical analyses were performed using SPSS statistical software (SPSS, Chicago, IL, USA). SHEsis was used to perform the Hardy–Weinberg equilibrium test and to compare differences in allele, genotype and haplotype frequencies between individuals with altered glucose metabolism and controls [6]. The population attributable risk was calculated as $(X-1)/X$, assuming the multiplicative model where $X = (1-f)^2 + 2f(1-f)\gamma + f^2\gamma^2; \gamma$ is the estimated OR and $f$ is the frequency of the risk allele [7].

The genotype frequencies for the five SNPs are summarised in Table 2; all were in Hardy–Weinberg equilibrium in the control group. The ORs of the three SNPs in *CDKAL1*, rs10946398 (OR 1.23, 95% CI 1.12–1.34, $p=1.49\times10^{-5}$), rs7754840 (OR 1.26, 95% CI 1.15–1.38, $p=5.44\times10^{-7}$) and rs7756992 (OR=1.16, 95% CI 1.06–1.27, $p=0.002$), were significantly associated with type 2 diabetes in the Han Chinese population. However, the ORs (1.16–1.26) were slightly higher than those identified in studies of European populations (1.12–1.20) [1–4], which implies some ethnic differences between these studies and our own. The SNP rs7754840 showed the most significant association with diabetes in our study. When we tested the association between *CDKAL1* rs7754840 and type 2 diabetes in a range of genetic models for our population, the most significant results were obtained in a recessive model that included age, sex and $\log_e$-BMI (CC vs GG/CG; OR 1.49, 95% CI 1.25–1.77, $p=7.4\times10^{-6}$). These results are consistent with studies on individuals of European descent and individuals from Hong Kong of Han Chinese ancestry [4], which showed that the risk for individuals homozygous for the rs7756992 risk allele is much higher than for those who are heterozygous. Assuming a population frequency of 39.6% for the risk allele, C, as observed in our controls, the population attributable risk was 17.8% for rs7754840 in the Han Chinese population.

In our study, the SNPs rs4402960 and rs1470579 in *IGF2BP2* were not associated with diabetes, but the ORs for the risk alleles were in the same direction as those in the Steinthorsdottir studies (rs4402960, T allele, OR 1.08; rs1470579, C allele, OR 1.03). We performed power calculations on the G*Power program [8], which has a 77% power to detect an OR of 1.14 (according to the OR in Europeans, as indicated by the Diabetes Genetics Initiative [DGI], Wellcome Trust Case Control Consortium [WTCCC] and Finland–United States Investigation of NIDDM Genetics [FUSION] [1]), given a frequency of a risk allele of 25% (rs4402960 in *IGF2BP2*) and a $p$ value of <0.01, so the absence of significant associations in our study may be due to a lack of power, mainly as a result of the limited sample number. It is important to note that the discrepancy might also be due to population-specific genetic backgrounds and environmental risk profiles.

Because the sample size of the IFG/IGT group was quite small we did not conduct a case–control comparison between this group and the control group, but the frequency of the risk allele in the IFG/IGT group was closer to that in the type 2 diabetic patient group than the control group.

In addition, we examined the association of these five SNPs with BMI, fasting plasma glucose, triacylglycerol, $HbA_{1c}$ and WHR in the control group using multiple linear regression with BMI, fasting plasma glucose, triacylglycerol, $HbA_{1c}$ and WHR as the dependent variables, genotype as the independent variable and sex, age and $\log_e$-BMI (except

**Table 2** Association study of type 2 diabetes susceptibility variants in the Han Chinese population

| Region, SNP | Risk allele | Genotype distribution | | | | Risk allele frequencies | | | Unadjusted[a] | | Adjusted[b] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Control | IFG/IGT | Type 2 diabetes | Control | IFG/IGT | Type 2 diabetes | Allele frequency | Additive model | Dominant model | Recessive model |
| *CDKAL1* | | | | | | | | | | | | |
| rs10946398 | C | CC | 293 (15.4) | 46 (21.2) | 372 (20.4) | 1,489 (39.1) | 191 (44.0) | 1,606 (44.1) | 1.23 (1.12–1.34) | 1.25 (1.13–1.38) | 1.28 (1.11–1.48) | 1.45 (1.21–1.74) |
| | | CA | 903 (47.5) | 99 (45.6) | 862 (47.3) | | | | $p_{freq}=1.49\times10^{-5}$ | $p_{add}=8.27\times10^{-6}$ | $p_{dom}=0.0008$ | $p_{rec}=5.16\times10^{-5}$ |
| | | AA | 707 (37.2) | 72 (33.2) | 588 (32.3) | | | | | | | |
| rs7754840 | C | CC | 314 (16.0) | 47 (21.3) | 402 (21.9) | 1,551 (39.6) | 190 (43.0) | 1,666 (45.3) | 1.26 (1.15–1.38) | 1.27 (1.16–1.40) | 1.31 (1.14–1.52) | 1.49 (1.25–1.77) |
| | | CG | 923 (47.2) | 96 (43.4) | 862 (46.9) | | | | $p_{freq}=5.44\times10^{-7}$ | $p_{add}=7.4\times10^{-7}$ | $p_{dom}=1.9\times10^{-4}$ | $p_{rec}=7.4\times10^{-6}$ |
| | | GG | 720 (36.8) | 78 (35.3) | 574 (31.2) | | | | | | | |
| rs7756992 | G | GG | 457 (24.3) | 67 (33.7) | 506 (29.8) | 1,870 (49.6) | 218 (54.8) | 1,812 (53.3) | 1.16 (1.06–1.27) | 1.18 (1.07–1.30) | 1.14 (0.97–1.34) | 1.36 (1.17–1.60) |
| | | GA | 956 (50.7) | 84 (42.2) | 800 (47.1) | | | | $p_{freq}=0.002$ | $p_{add}=0.0008$ | $p_{dom}=0.11$ | $p_{rec}=0.0001$ |
| | | AA | 471 (25.0) | 48 (24.1) | 394 (23.2) | | | | | | | |
| *IGF2BP2* | | | | | | | | | | | | |
| rs4402960 | T | TT | 128 (6.5) | 13 (5.6) | 141 (7.6) | 982 (24.9) | 117 (25.2) | 982 (26.4) | 1.08 (0.97–1.19) | 1.06 (0.95–1.18) | 1.07 (0.93–1.22) | 1.11 (0.86–1.45) |
| | | TG | 726 (36.9) | 91 (39.2) | 700 (37.6) | | | | $p_{freq}=0.15$ | $p_{add}=0.29$ | $p_{dom}=0.36$ | $p_{rec}=0.42$ |
| | | GG | 1,116 (56.6) | 128 (55.2) | 1,022 (54.9) | | | | | | | |
| rs1470579 | C | CC | 124 (7.5) | 16 (7.0) | 148 (7.9) | 902 (27.2) | 122 (26.6) | 1,043 (27.7) | 1.03 (0.93–1.14) | 1.02 (0.92–1.14) | 1.03 (0.89–1.18) | 1.03 (0.80–1.34) |
| | | CA | 654 (39.4) | 90 (39.3) | 747 (39.7) | | | | $p_{freq}=0.60$ | $p_{add}=0.68$ | $p_{dom}=0.69$ | $p_{rec}=0.82$ |
| | | AA | 883 (53.2) | 123 (53.7) | 987 (52.4) | | | | | | | |

Data are presented as *n* (%) or OR (95% CI)

[a] Differences in allele frequencies (*p*freq) not adjusted for age, sex or $\log_e$-BMI; *p*freq values were calculated using Fisher's exact test

[b] $p_{add}$ (additive model), *p*dom (dominant model) and *p*rec (recessive model) were assessed by logistic regression after adjusting for sex, age and $\log_e$-BMI

where BMI was the dependent variable) as covariates (data not show). The risk alleles of rs10946398 and rs7754840 showed significant associations with fasting plasma glucose ($p=1.1\times10^{-3}$ and $p=2.1\times10^{-3}$, respectively), and the associations were much stronger for genotypes homozygous for the risk alleles. The SNPs rs4402960 and rs1470579 in the *IGF2BP2* region were associated with WHR ($p=0.02$ and $p=0.03$, respectively), but neither SNP showed an association with BMI, indicating that *IGF2BP2* may affect adipose distribution.

In conclusion, we have identified a positive association between SNPs within *CDKAL1* in the Han Chinese population. More comprehensive studies are needed into the associations of variations in *CDKAL1* with functional correlates and type 2 diabetes risk in larger, ethnically diverse populations.

**Acknowledgements** We appreciate the contribution of all participants in this study, as well as that of the doctors who helped us in the diagnosis. This work was supported by the National Key Technology R&D Program (2006BAI05A05), the Chinese Nutrition Society (05015), the Dannon Institute, the Shanghai-Unilever Research and Development Fund (06SU07007), the Shanghai Municipality Science & Technology Commission (05JC14090), the Shanghai Leading Academic Discipline Project (B205) and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-YW-R-01).

**Duality of interest** The authors declare that there is no duality of interest associated with this manuscript.

## References

1. Saxena R, Voight BF, Lyssenko V et al (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316:1331–1336
2. Zeggini E, Weedon MN, Lindgren CM et al (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 316:1336–1341
3. Scott LJ, Mohlke KL, Bonnycastle LL et al (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316:1341–1345
4. Steinthorsdottir V, Thorleifsson G, Reynisdottir I et al (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 39:770–775
5. Omori S, Tanaka Y, Takahashi A et al (2007) Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8 and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. Diabetes 57:791–795
6. Shi YY, He L (2005) SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. Cell Res 15:97–98
7. Altshuler D, Hirschhorn JN, Klannemark M et al (2000) The common PPARγ(Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 26:76–80
8. Faul F, Erdfelder E, Lang AG, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 39:175–191