Association of A Common Haplotype of Hepatocyte Nuclear Factor 1α With Type 2 Diabetes in Chinese Population[1]

CONG-RONG WANG, CHENG HU, RONG ZHANG, QI-CHEN FANG, XIAO-JING MA, WEI-PING JIA, AND KUN-SAN XIANG[2]

Shanghai Diabetes Institute, Department of Endocrinology and Metabolism, Shanghai Jiaotong University Affiliated Sixth People's Hospital, Shanghai 200233,China

Objective To analyze the association of variants of hepatocyte nuclear factor-1α (*HNF-1α*) gene with type 2 diabetes in Chinese population. **Methods** In 152 unrelated type 2 diabetes patients and 93 unrelated controls, eleven single nucleotide polymorphisms (SNPs) were identified and genotyped. Statistical analyses were performed to investigate whether these SNPs were associated with diabetes status in our samples. **Results** In the individual SNP study, no SNP differed significantly in frequency between type 2 diabetes patients and controls. In the haplotype analysis, two haplotype blocks were identified. In haplotype block 1, no evidence was found between common *HNF-1α* haplotypes and type 2 diabetes. However, in haplotype block 2, a common haplotype GCGC formed by four tagging SNPs (tSNPs) was found to be associated with decreased risk of type 2 diabetes (odds ratio [OR] 0.6011, 95% confidence interval [CI] 0.4138-0.8732, *P*=0.0073, empirical *P*=0.0511, permutation test). A similar trend was also observed in the diplotype analysis, indicating that the increasing copy number of the haplotype GCGC was associated with the decreased frequency of diabetes (*P*=0.0193). **Conclusion** The results of this study provide evidence that the haplotype of *HNF-1α* decreases the risk of type 2 diabetes in Chinese individuals.

Key words: Hepatocyte nuclear factor-1α; Type2 diabetes; SNP; Haplotype analysis

INTRODUCTION

Hepatocyte nuclear factor-1α (*HNF-1α*, also known as TCF1) is a homeodomain-containing transcription factor which was originally identified in liver. *HNF-1α* gene is expressed in tissues of liver, kidney, intestine, pancreas, and plays an important role in the development and differentiation of these organs[1]. *HNF-1α*, together with other transcription factors, constitutes a complicated network that controls organ-specific gene expression during embryonic development and in adult tissues.

Mutations of the *HNF-1α* gene lead to β-cell dysfunction and maturity-onset of diabetes in the young (MODY 3). In the British population, *HNF-1α* mutations account for nearly two-thirds of all MODY families, while they are not a major cause of early-onset and/or multiply affected diabetic pedigrees in Chinese population[2-3]. Meanwhile, some common amino acid variants of *HNF-1α* have been identified to be associated with type 2 diabetes[4-6] and related intermediate traits[7-12]. Recently, Jackson *et al*.[13] have also found that the variant A98V of *HNF-1α*, together with other variants of *NEUROD1* and *NEUROG3*, has a combined effect on glucose tolerance in the South Indian population. These studies suggest that *HNF-1α* might play a role in the pathogenesis of common type 2 diabetes.

Linkage disequilibrium (LD) and haplotype analysis have been used to test the association of single nucleotide polymorphisms (SNPs) of candidate genes with type 2 diabetes within the case-control population in recent studies[14-19]. The new statistical

[1]This research was supported by the National 973 Program (2006 CB503901), the Major Project of Shanghai Science and Technology Development Foundation (02DJ14052-I), the Science and Technology Development Foundation of Shanghai Municipal Health Bureau (034043) and the Young Doctor Training Project of Shanghai Municipal Health Bureau.

[2]Correspondence should be addressed to Professor Kun-San XIANG, Shanghai Diabetes Institute, Shanghai Jiaotong University Affiliated Sixth People's Hospital, 600 Yishan Rd, Shanghai 200233, China. Tel: 86-21-64369181-8720. Fax: 86-21-64368031. E-mail: sphxiang@163.com

Biographical note of the first author: Cong-Rong WANG, born in 1972, M. D., Attending Physician, majoring in genetics of diabetes. E-mail: wcr601@hotmail.com

0895-3988/2007
CN 11-2816/Q

methods provide more information than before. Even if no individual SNPs are associated with disease status, it may be possible to identify haplotypes that are significantly associated with increased or decreased disease risk.

The promoter, coding and flanking intron regions of *HNF-1α* was scanned in our previous study and several variants have been identified[2]. In this study, analyses of the variants were performed to assess their association and potential role in type 2 diabetes in Chinese population.

SUBJECTS AND METHODS

Two hundred and forty-five unrelated Chinese residents in Shanghai were selected in this case-control study. The type 2 diabetes patients included one hundred and fifty-two unrelated probands from early-onset and/or multiplex diabetes families. Glucose tolerance status was identified according to the 1999 WHO criteria[20]. Diabetic patients who had antibodies to tyrosine phosphatase-like protein (IA2-Ab) or GAD were excluded. Mutation carriers of MODY genes (MODY1-6) or mitochondrial nucleotide 3243 A-to-G substitution were also excluded by genetic testing. The age of diabetes patients at diagnosis was 46.41±10.57 years, and BMI was 24.08±3.30 kg/m^2. Ninety-three unrelated controls were selected from a community-based survey for type 2 diabetes. All the controls were old "super normal" individuals (age > 65 years, BMI < 25 kg/m^2) without hyperglycemia, hypertension or dyslipidaemia[20]. These controls had no family history of diabetes. This study was approved by the Institutional Review Board of Shanghai Jiaotong University Affiliated Sixth People's Hospital. Informed consents were obtained from all the participants.

Genotyping

The promoter, coding regions, and flanking intron regions of *HNF-1α* gene were screened by PCR-SSCP and sequencing analysis. Eleven variants were identified in the initial research[2] which covered nearly 22kb of *HNF-1α* gene. These SNPs were typed by direct sequencing or PCR-RFLP in all the diabetes patients and controls. The sequencing reactions were analyzed with the Applied Biosystems 3100 Genetic Analyzer in both directions. The primers were described in a previous study[2].

Statistical Analysis

Allele frequencies were determined by gene counting, and observed genotypes were tested for fit to the expectations of Hardy-Weinberg equilibrium using the χ^2 test.

SNPs with a minor allele frequency (MAF) over 1% were selected for analysis. Pairwise linkage disequilibrium was calculated from the combined data of patients and controls by using |D'| and r^2. Haplotype blocks were defined according to the confidence interval algorithm[21] and haplotype frequencies were estimated for the combined set of patients and controls using the expectation-maximization algorithm and performed in Haploview (v 3.2)[22]. Tagging SNPs (tSNPs) were chosen for further analysis. Diplotype of each subject was inferred by PHASE (v 2.1)[23-24]. Differences in SNP, haplotype and diplotype distribution between patients and controls were tested using χ^2 test. Odds ratios (ORs) and *P* values were determined by SAS for WINDOWS (v 6.12).

Permutation tests were implemented by Haploview (v 3.2)[22] to estimate the probability of a similar number of associated individual SNPs or haplotypes due to chance alone. Ten thousand permuted samples were created by randomly shuffling genotypes. The overall *P* value was computed as the proportion of permuted samples having at least individual SNPs or halplotypes with *P* values less than or equal to the least significant *P* value in the original data.

RESULTS

Association Analysis of Individual SNPs

Nine common variants with MAF >1% were used in the present study. All SNPs were in Hardy-Weinberg equilibrium. The allele frequency of each SNP did not differ significantly between type 2 diabetes patients and controls (Table 1).

Association Analysis of Haplotype/diplotype

Table 2 shows the pairwise LD (|D'| and r^2) among the genotyped SNPs in Chinese subjects. Two haplotype blocks were observed in the *HNF-1α* region in the studied subjects (Fig. 1). In block 1, two tSNPs (rs1169289 and rs1169288) that defined three haplotypes with a frequency >5%, accounted for 99.6% of all haplotypes across the *HNF-1α* region, while in block 2, four tSNPs (rs1169294, rs3751156, rs1169301, rs2259820) that defined six haplotypes with a frequency >5%, accounted for 93.8% of all haplotypes across the region (Table 3).

In haplotype block 1, no haplotype differed significantly in frequency between patients and contros (Table 3A), while in haplotype block 2, a significant difference was found in the frequency of

TABLE 1

Type 2 Diabetes Association Analysis of Nine Individual SNPs in Patients and Controls

SNP Name	[a]SNP Position (bp)	Major/Minor Alleles	Minor Allele Frequency		Case-control Chi-square	*P* Value
			Patients	Controls		
rs1169289	119879342	C/G	0.405	0.344	1.791	0.1808
rs1169288	119879370	A/C	0.454	0.382	2.461	0.1167
rs1169294	119889314	G/A	0.414	0.344	2.408	0.1207
rs1169301	119894020	C/T	0.477	0.419	1.545	0.2139
rs3751156	119897022	G/T	0.141	0.108	1.185	0.2763
rs2259820	119898062	C/T	0.487	0.419	2.115	0.1459
rs2464196	119898147	G/A	0.477	0.419	1.545	0.2139
rs2464195	119898195	G/A	0.473	0.419	1.375	0.2410
rs735396	119901564	T/C	0.484	0.419	1.915	0.1664

Note. [a]SNP positions are indicated in base pair from the p-terminus of chromosome 12, Haplotype tagging SNPs are shaded.

TABLE 2

Pairwise Linkage Disequilibrium Among *HNF-1α* SNPs in the Studied Subjects

SNP Name	rs1169289	rs1169288	rs1169294	rs1169301	rs3751156	rs2259820	rs2464196	rs2464195	rs735396
rs1169289	--	0.80	0.73	0.53	0.05	0.57	0.60	0.60	0.62
rs1169288	0.98	--	0.62	0.47	0.02	0.47	0.51	0.50	0.52
rs1169294	0.86	0.85	--	0.62	0.09	0.66	0.68	0.69	0.70
rs1169301	0.84	0.72	0.90	--	0.11	0.79	0.82	0.82	0.84
rs3751156	0.74	0.41	1.00	0.94	--	0.10	0.10	0.10	0.10
rs2259820	0.89	0.73	0.94	0.89	0.89	--	0.96	0.97	0.94
rs2464196	0.90	0.75	0.94	0.90	0.89	0.99	--	0.99	0.98
rs2464195	0.89	0.74	0.94	0.90	0.89	1.00	1.00	--	0.98
rs735396	0.92	0.77	0.96	0.92	0.89	0.97	1.00	1.00	--

Note. |D'| values are shown in the lower left, while r^2 values are shown in the upper right.

TABLE 3A

Association Analysis of Haplotype Frequencies and Type 2 Diabetes in Haplotype Block 1

rs1169289	rs1169288	Frequency		Chi-square	OR (95%CI)	[†]*P* Value	[‡]Empirical *P* Value
C/G	A/C	Case	Control				
C	A	0.539	0.618	2.921	0.7223(0.4979-1.0477)	0.0874	0.5550
G	C	0.398	0.344	1.430	1.2587(0.8609-1.8402)	0.2319	0.9139
C	C	0.056	0.038	0.823	1.5241(0.6203-3.7452)	0.3642	0.9960

TABLE 3B

Association Analysis of Haplotype Frequencies and Type 2 Diabetes in Haplotype Block 2

Haplotype	rs1169294	rs1169301	rs3751156	rs2259820	Frequency		Chi-square	[*]OR (95%CI)	[†]*P* Value	[‡]Empirical *P* Value
	G/A	C/T	G/T	C/T	Case	Control				
ht1	G	C	G	C	0.340	0.461	7.192	0.6011(0.4138-0.8732)	0.0073	0.0511
ht2	A	T	G	T	0.384	0.333	1.334	1.2509(0.8532-1.8340)	0.2481	0.9206
ht3	G	C	T	C	0.130	0.107	0.539	1.2359(0.6975-2.1899)	0.4628	0.9995
ht4	G	T	G	T	0.053	0.086	2.102	0.5942(0.2900-1.2175)	0.1427	0.8111

Note. Haplotypes with estimated frequencies >5% are shown. [*]OR: odds ratio. CI: confidence interval. [†]*P* values <0.05 are shown in bold. [‡]Empirical *P* values are obtained from 10 000 permutation test.

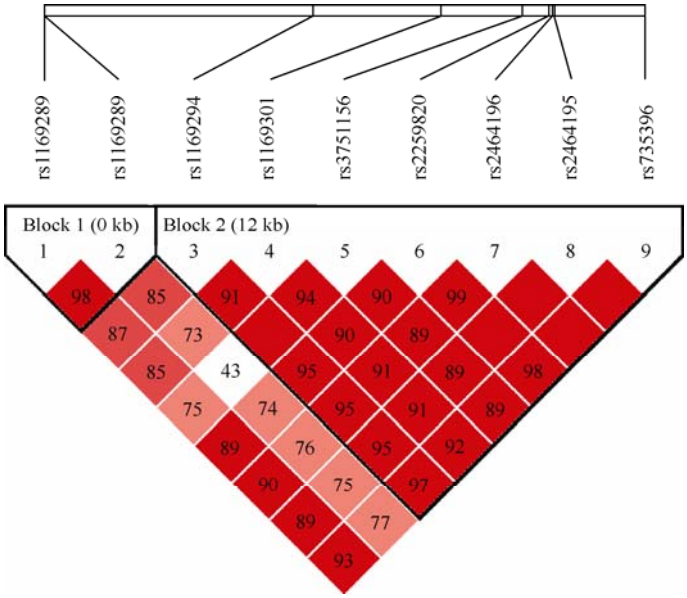

FIG. 1. Linkage disequilibrium (LD) plot and block structure of *HNF-1α* region in the studied subjects. The block structure based on confidence intervals |D' | is shown as triangles. Each diamond represents the pairwise magnitude of LD, with red indicating strong LD (|D'| > 0.8) and a logarithm of odds score ⩾2.0.

TABLE 4

Type 2 Diabetes Association Analysis of GCGC Haplotype Combination (Diplotype)

Affection Status	Diplotype Distribution, *n* (%)			
	ht1/ht1	ht1/-	-/-	*P* Value
Case	18(11.84%)	67(44.08%)	67(44.08%)	0.0193
Control	18(19.35%)	50(53.67%)	25(26.88%)	χ^2=7.8936

Note. ht1: GCGC haplotype.

the most common haplotype GCGC (ht1) between patients and controls (empirical *P*=0.0511, permutation test). The frequency of ht1 haplotype in the patients with type 2 diabetes was significantly lower than that in the "super normal" controls (34.0% *vs* 46.1%). The patients had an estimated OR 0.6011 (*P*=0.0073, 95%CI=0.4979-1.0477), indicating a significant decrease of relative risk of type 2 diabetes with this haplotype (Table 3B). Since the most common haplotype GCGC (ht1) seemed to be the protective haplotype, the haplotype combination (diplotype) analysis was then performed. The increasing copy number of the ht1 was found to be associated with a lower frequency of type 2 diabetes (*P*=0.0193) (Table 4).

DISCUSSION

In the present work, variants in the *HNF-1α* gene promoter, coding and flanking introns were delineated in type 2 diabetes patients and controls. Two haplotype blocks containing three and six common haplotypes respectively, were observed.

Using haplotype/diplotype analyses, a common haplotype identified in the second block of the *HNF-1α* gene was found to be associated with disease status. The GCGC haplotype, formed by rs1169294, rs3751156, rs1169301, and rs2259820, was associated with a low risk of type 2 diabetes (OR=0.6011). Moreover, permutation tests indicated that the observed association was unlikely to occur by chance (empirical *P*=0.0511). The GCGC haplotype covers the regions that encode DNA-binding domain and transactivation domain of *HNF-1α* gene, which is essential to the gene function. The significantly different haplotype distributions in block 2 between patients with type 2 diabetes and controls suggest that the *HNF-1α* gene locus contributes to the susceptibility to type 2 diabetes. In this study, no individual SNP was associated with type 2 diabetes.

One of the possible reasons is that the effects of individual SNPs are relatively weak. On the other hand, it is also possible that the GCGC haplotype tags an additional locus not genotyped.

Although mutations *HNF-1α* are well recognized to be one of the causes of MODY phenotype, recent studies have shown that coding or noncoding variants of MODY genes may be associated with type 2 diabetes. At least three individual SNPs of the *HNF-1α* gene have been reported to be associated with type 2 diabetes and related traits in some populations. For example, the Ala98Val variant is associated with type 2 diabetes and decreased insulin secretion during OGTT in Finnish[5] and Danish subjects, respectively[8,11]. The G319S and I27L variants are associated with lipid levels in Canadian Oji-cree subjects and Japanese, respectively[7,10]. In Chinese subjects, only I27L (rs1169288) variant has been identified, which has no association with type 2 diabetes in individual SNP analysis. In particular, there is evidence that the haplotype containing I27L does not predispose to type 2 diabetes. However, analyses to identify diabetes-related quantitative traits remain necessary in future studies.

To our knowledge, the present study is the first to report that the haplotype in *HNF-1α* gene may affect the low risk of type 2 diabetes. Although two large independent Caucasian studies showed that no common haplotypes of *HNF-1α* are associated with type 2 diabetes[25-26], our findings are scientifically justified because patients were enrolled from early onset and/or multiplex diabetes families, which means that they may be enriched for genetic effects. Moreover, the normal glucose status of our controls was confirmed after the age of sixty-five years, which means that the controls unlikely develop diabetes. However, the role of *HNF-1α* haplotypes remains to be determined in large samples of our population.

In conclusion, the variant combinations in the *HNF-1α* gene are associated with a low-risk of type 2 diabetes in Chinese population. Further studies in other populations and functional analysis, are required to elucidate the role of variation in *HNF-1α* in the pathogenesis of type 2 diabetes.

ACKNOWLEDGEMENTS

We acknowledge Maggie C. Y. Ng, from the Department of Medical and Therapeutics, the Chinese University of Hong Kong, for her excellent suggestions for the genetic analysis. We also thank Jun-Xi LU, Jing XU, Jun-Ling TANG, Xiao-Ping PAN, and Qiao-Ying YUAN for their technical supports.

REFERENCES

1. Pontoglio M, Barra J, Hadchouel M, *et al.* (1996). Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome. *Cell* **84**(4), 575-585.
2. Fang Q C, Zhang R, Wang C R, *et al.* (2004). Scanning *HNF-1α* gene mutation in Chinese early-onset and/or multiplex diabetes pedigrees. *Zhong Hua Yi Xue Yi Chuan Xue Za Zhi* **21**(4), 329-334.
3. Frayling T M, Evans J C, Bulman M P, *et al.* (2001).Beta-cell genes and diabetes: molecular and clinical characterization of mutations in transcription factors. *Diabetes* **50**(Suppl 1), 94-100.
4. Hegele R A, Cao H, Harris S B, *et al.* (1999). The hepatic nuclear factor-1alpha G319S variant is associated with early-onset type 2 diabetes in Canadian Oji-Cree. *J Clin Endocrinol Metab* **84**(3), 1077-1082.
5. Rissanen J, Wang H, Miettinen R, *et al.* (2000). Variants in the hepatocyte nuclear factor-1alpha and -4alpha genes in Finnish and Chinese subjects with late-onset type 2 diabetes. *Diabetes Care* **23**, 1533-1538.
6. Anruadha S, Radha V, Deepa R, *et al.* (2005). A Prevalent Amino Acid Polymorphism at Codon 98 (Ala98Val) of the Hepatocyte Nuclear Factor-1α Is Associated With Maturity-Onset Diabetes of the Young and Younger Age at Onset of Type 2 Diabetes in Asian Indians. *Diabetes Care* **28**, 2430-2435.
7. Hegele R A, Cao H N, Harris S B, *et al.* (2000). The Private Hepatocyte Nuclear Factor-1 α G319S Variant Is Associated With Plasma Lipoprotein Variation in Canadian Oji-Cree. *Arterioscler. Thromb Vasc Biol* **20**(1), 217-222.
8. Urhammer S A, Hansen T, Ekstrom C T, *et al.* (1998). The Ala/Val98 Polymorphism of the Hepatocyte Nuclear Factor-1α Gene Contributes to the Interindividual Variation in Serum C-Peptide Response during an Oral Glucose Tolerance Test: Evidence from Studies of 231 Glucose-Tolerant First Degree Relatives of Type 2 Diabetic Probands. *J Clin Endocrinol Metab* **83**(12), 4506-4509.
9. Chiu K C, Chuang L, Ryu J M, *et al.* (2000). The I27L Amino Acid Polymorphism of Hepatic Nuclear Factor-1 α Is Associated with Insulin Resistance. *J Clin Endocrinol Metab* **85**(6), 2178-2183.
10. Babaya N, Ikegami H, Fujisawa T, *et al.* (2003). Association of I27L Polymorphism of Hepatocyte Nuclear Factor-1α Gene with High-Density Lipoprotein Cholesterol Level. *J Clin Endocrinol Metab* **88**(6), 2548-2551.
11. Urhammer S A, Fridberg M, Hansen T, *et al.* (1997). A prevalent amino acid polymorphism at codon 98 in the hepatocyte nuclear factor-1alpha gene is associated with reduced serum C-peptide and insulin responses to an oral glucose challenge. *Diabetes* **46**, 912-916.
12. Lauenborg J, Damm P, E K J, *et al.* (2004). Studies of the Ala/Val98 polymorphism of the hepatocyte nuclear factor-1a gene and the relationship to ß-cell function during an OGTT in glucose-tolerant women with and without previous gestational diabetes mellitus. *Diabet Med* **21**, 1310-1315.
13. Jack A E, Cassell P G, North B V, *et al.* (2004). Polymorphic Variations in the Neurogenic Differentiation-1, Neurogenin-3, and Hepatocyte Nuclear Factor-1α Genes Contribute to Glucose Intolerance in a South Indian Population. *Diabetes* **53**, 2122-2125.
14. Hu F B, Doria A, Li T, *et al.* (2004). Genetic Variation at the Adiponectin Locus and Risk of Type 2 Diabetes in Women. *Diabetes* **53**, 209-213.
15. Florez J C, Burtt N, Bakker P I W, *et al.* (2004). Haplotype Structure and Genotype-Phenotype Correlations of the Sulfonylurea Receptor and the Islet ATP-Sensitive Potassium Channel Gene Region. *Diabetes* **53**, 1360-1368.

16. Oberkofler H, Linnemayr V, Weitgasser R, *et al.* (2004). Complex Haplotypes of the PGC-1α Gene Are Associated With Carbohydrate Metabolism and Type 2 Diabetes. *Diabetes* **53**, 1385-1393.
17. Flavell D M, Ireland H, Stephens J W, *et al.* (2005). Peroxisome Proliferator-Activated Receptor α Gene Variation Influences Age of Onset and Progression of Type 2 Diabetes. *Diabetes* **54**, 582-586.
18. Song Y Q, Niu T H, Manson J E, *et al.* (2004). Are Variants in the CAPN10 Gene Related to Risk of Type 2 Diabetes? A Quantitative Assessment of Population and Family-Based Association Studies. *Am J Hum Genet* **74**, 208-222.
19. Silander K, Mohlke K L, Scott L J, *et al.* (2004). Genetic Variation Near the Hepatocyte Nuclear Factor-4α Gene Predicts Susceptibility to Type 2 Diabetes. *Diabetes* **53**, 1141-1149.
20. World Health Organization (1999). Definition, diagnosis and classification of diabetes mellitus and its complications: report of a WHO consultation. Part 1: Diagnosis and classification of diabetes mellitus. WHO, Geneva.
21. Gabriel S B, Schaffner S F, Nguyen H, *et al.* (2002). The structure of haplotype blocks in human genome. *Science* **296**(5576), 2225-2229.
22. Barrett J C, Fry B, Maller J, *et al.* (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265.
23. Stephens M, Scheet P (2005). Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *American J Hum Genet* **76**, 449-462.
24. Stephens M, Smith N J, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *American J Hum Genet* **68**, 978-989.
25. Winckler W, Burtt N P, Holmkvist J, *et al.* (2005). Association of common variation in the *HNF1α* gene region with risk of type 2 diabetes. *Diabetes* **54**, 2336-2342.
26. Weedon M N, Owen K R, Shields B, *et al.* (2005). A large-scale association analysis of common variation of the *HNF1α* gene with type 2 diabetes in the U.K. Caucasian population. *Diabetes* **54**, 2487-2491.

(*Received January* 9, 2006 *Accepted December* 2, 2006)