Association Study of the Genetic Polymorphisms of the Transcription Factor 7-Like 2 (*TCF7L2*) Gene and Type 2 Diabetes in the Chinese Population

Yi-Cheng Chang,[1] Tien-Jyun Chang,[2] Yi-Der Jiang,[2] Shan-Shan Kuo,[2] Kuan-Ching Lee,[2] Ken C. Chiu,[3] and Lee-Ming Chuang[2,4]

OBJECTIVE—Genetic polymorphisms of the transcription factor 7-like 2 (*TCF7L2*) gene is one of the few validated genetic variants with large effects on the risk of type 2 diabetes in the populations of European ancestry. In this study, we aimed to explore the effect of the *TCF7L2* polymorphisms in a Han Chinese population.

RESEARCH DESIGN AND METHODS—We genotyped 20 single nucleotide polymorphisms (SNPs) across the *TCF7L2* gene in 1,520 unrelated subjects from a Han Chinese population in Taiwan. The associations of SNPs and haplotypes with type 2 diabetes and linkage disequilibrium (LD) structure of the *TCF7L2* gene were analyzed.

RESULTS—The previously reported SNPs rs7903146 T- and rs12255372 T-alleles of the *TCF7L2* gene were rare and were not associated with type 2 diabetes in a Chinese population, which may attribute to the low frequencies of these two SNPs. SNP rs290487 located in an LD block close to the 3′ end of the gene was associated with type 2 diabetes (allele-specific $P = 0.0021$; permuted $P = 0.03$). The odds ratio was 1.36 for the CT genotype (95% CI 1.08−1.71; $P = 0.0063$) and 1.51 for the CC genotype (1.10 −2.07; $P = 0.0085$) compared with the TT genotype, corresponding to a population attributable risk fraction of 18.7%. The haplotypes composed of rs290487 were also significantly associated with type 2 diabetes (global $P = 0.012$).

CONCLUSIONS—We identified a novel risk-conferring genetic variant of *TCF7L2* for type 2 diabetes in a Chinese population. Our data suggested that the *TCF7L2* genetic polymorphisms are major determinants for risk of type 2 diabetes in the Chinese population. ***Diabetes*** **56:2631–2637, 2007**

Type 2 diabetes is a highly inheritable metabolic disorder of polygenetic nature (1). Although theoretical analyses emphasized the power of genetic association study in common multifactorial diseases, the search for genes that increase the risk of type 2 diabetes has not been very successful so far. The genes implicated in type 2 diabetes confer only modest effects on the disease risk and in many cases have yielded inconsistent results in replication efforts (2). Only few associations, notably the Pro12Ala polymorphism in the peroxisome proliferator–activated receptor (*PPARG*) gene (3), the Glu23Lys polymorphism in the *KCNJ11* gene (4), and the genetic variants of calpain-10 genes (5), have been convincingly replicated.

Recently, researchers seeking the cause of a previously identified linkage signal on chromosome 10q found a strong association of a common microsatellite (DG10S478) of the transcription factor 7-like 2 gene (*TCF7L2*) with type 2 diabetes in an Icelandic sample, and the result was replicated in the samples from the U.S. and Denmark (6). DG10S478 is located within a well-defined linkage disequilibrium (LD) block of 92.1 kb that encompassed exon 4 and parts of two large flanking introns. Five single nucleotide polymorphisms (SNPs) (rs12255372, rs7903146, rs7901695, rs11196205, and rs7895340) within the LD block also showed similarly robust associations with type 2 diabetes (6). Further studies in other European populations, African Americans, Mexican Americans, and Asian Indians confirmed the strong associations with an estimated population attributable risk of 17–28% (7–22). Several genome-wide association studies independently confirmed the strong associations of SNP rs7903146 in the *TCF7L2* locus with type 2 diabetes (23–26). Further phylogenetic reconstruction of the revolutionary relationship between the haplotypes within the exon 4 LD block identified the $HapB_{T2D}$ lineage, which contains all haplotypes carrying the ancestral T-allele of SNP rs7903146 as a type 2 diabetes risk variant (27). These data convincingly demonstrated that the genetic variants within the exon 4 LD block of the *TCF7L2* gene, especially rs7903146 T, are major genetic risk factors for type 2 diabetes in populations of European ancestry.

However, unlike in populations of European origin, the frequency of the rs7903146 T-allele is relatively low (~2%) in the HapMap Han Chinese Beijing (CHB) databank, raising the question of whether these variants are major contributors of type 2 diabetes in the Chinese population. Furthermore, the association study of the genetic polymorphism of *TCF7L2* with type 2 diabetes in the Chinese population has been lacking. Therefore, we conducted an association study in 1,520 unrelated subjects from a Han Chinese population in Taiwan. SNPs tagging common variations across the *TCF7L2* gene were selected according the HapMap CHB databank group (http://www.hapmap.org/), and additional SNPs with previously known strong associations with type 2 diabetes in populations of Euro-

From the [1]Department of Internal Medicine, National Taiwan University Hospital Yunlin Branch, Yunlin, Taiwan; the [2]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan; the [3]Department of Diabetes, Endocrinology, and Metabolism, City of Hope National Medical Center, Duarte, California; and the [4]Graduate Institute of Clinical Medicine, National Taiwan University Medical College, Taipei, Taiwan.

Address correspondence and reprint requests to Lee-Ming Chuang, MD, PhD, Department of Internal Medicine, National Taiwan University Hospital, 7 Chung-Shan South Rd., Taipei, Taiwan. E-mail: leeming@ntu.edu.tw.

Received for publication 27 March 2007 and accepted in revised form 14 June 2007.

Published ahead of print at http://diabetes.diabetesjournals.org on 19 June 2007. DOI: 10.2337/db07-0421.

LD, linkage disequilibrium; PAF, population attributable risk fraction; SNP, single nucleotide polymorphism.

pean ancestry were also genotyped. Their associations with type 2 diabetes and LD structure were analyzed.

RESEARCH DESIGN AND METHODS

The studied population comprised 760 normoglycemic control subjects and 760 type 2 diabetic patients from a Han Chinese population in Taiwan. The glucose-tolerant (determined by 75-g oral glucose tolerance tests) control subjects were recruited from routine health examinations, and type 2 diabetic patients were recruited from the metabolism clinics of the National Taiwan University Hospital. Type 2 diabetic patients were diagnosed based on the criteria of the American Diabetes Association (28) with fasting plasma glucose ≥126 mg/dl or 2-h plasma glucose ≥200 mg/dl during an oral glucose tolerance test. Patients with ages of onset <35 years were excluded. Total cholesterol and triglycerides were measured in fasting samples using an autoanalyzer (Hitachi 7250 special; Hitachi, Tokyo, Japan). A1C was measured by high-performance liquid chromatography (Primus, Wichita, KS). Written informed consent was obtained from every participating subject, and the study was approved by the institutional review board of the National Taiwan University Hospital.

Selection of tag SNPs and genotyping. From HapMap CHB databank (public data release 21 a/phase II, Jan. 2007), 112 SNPs that had minor allele frequencies exceeding 1% without violation of Hardy-Weinberg equilibrium were identified. To identify common haplotype tagging SNPs, these eligible SNPs were entered into the Tagger program implemented in Haploview version 3.32 (29) (http://www.broad.mit.edu/mpg/haploview/) with a minor allele frequency threshold of 20% and an r^2 of 0.8. Thirteen tag SNPs (rs6585194, rs7919409, rs11196219, rs4918792, rs10749127, rs11196224, rs7085532, rs17130188, rs10787475, rs12775879, rs290489, rs290487, and rs290481) were selected, which captured 92.2% of the variance for haplotypes composed of all SNPs with minor allele frequencies ≥20% across the *TCF7L2* gene. Because nearly all SNPs within the exon 4 LD block have minor allele frequencies <5% in the HapMap CHB databank, we additionally selected eight SNPs with known associations with type 2 diabetes (rs4506565, rs12243326, rs7895340, rs7904146, rs12255372, rs7079711, rs11196192, and rs11196213) within this region. Genotyping was performed using the GenomeLab SNPstream genotyping platform (Beckman Coulter) and its accompanying SNPstream software suite (30).

Statistical analyses. A Hardy-Weinberg equilibrium test was performed for each sequence variant for the control group before marker-trait association analysis. Tests for the associations of each SNP and haplotype with type 2 diabetes were estimated by using the Haploview software (29). Nominal two-sided *P* values were reported and were corrected for multiple testings by permutation 10,000 times. For haplotype construction, genotype data of both case and control groups were used to estimate intermarker LD by measures pairwise D′ and r^2 and define LD blocks. We used the solid spine of the LD method implemented in the Haploview software to define an LD block with an extended spine if D′ was >0.8 (29).

For the sliding window association analysis of the haplotypes composed of the adjacent 2–10 SNPs, we used the score test developed by Schaid et al. (31) and implemented in the Haplo.Stats package (http://www.mayo.edu/hsr/Sfunc.html). This method allows the adjustment for possible confounding variables and provides both global tests and haplotype-specific tests. The global *P* value for the associations of haplotypes composed of the adjacent 2–10 SNPs with type 2 diabetes were estimated with simulations 10,000 times. A *P* value <0.05 was considered statistically significant. We also performed

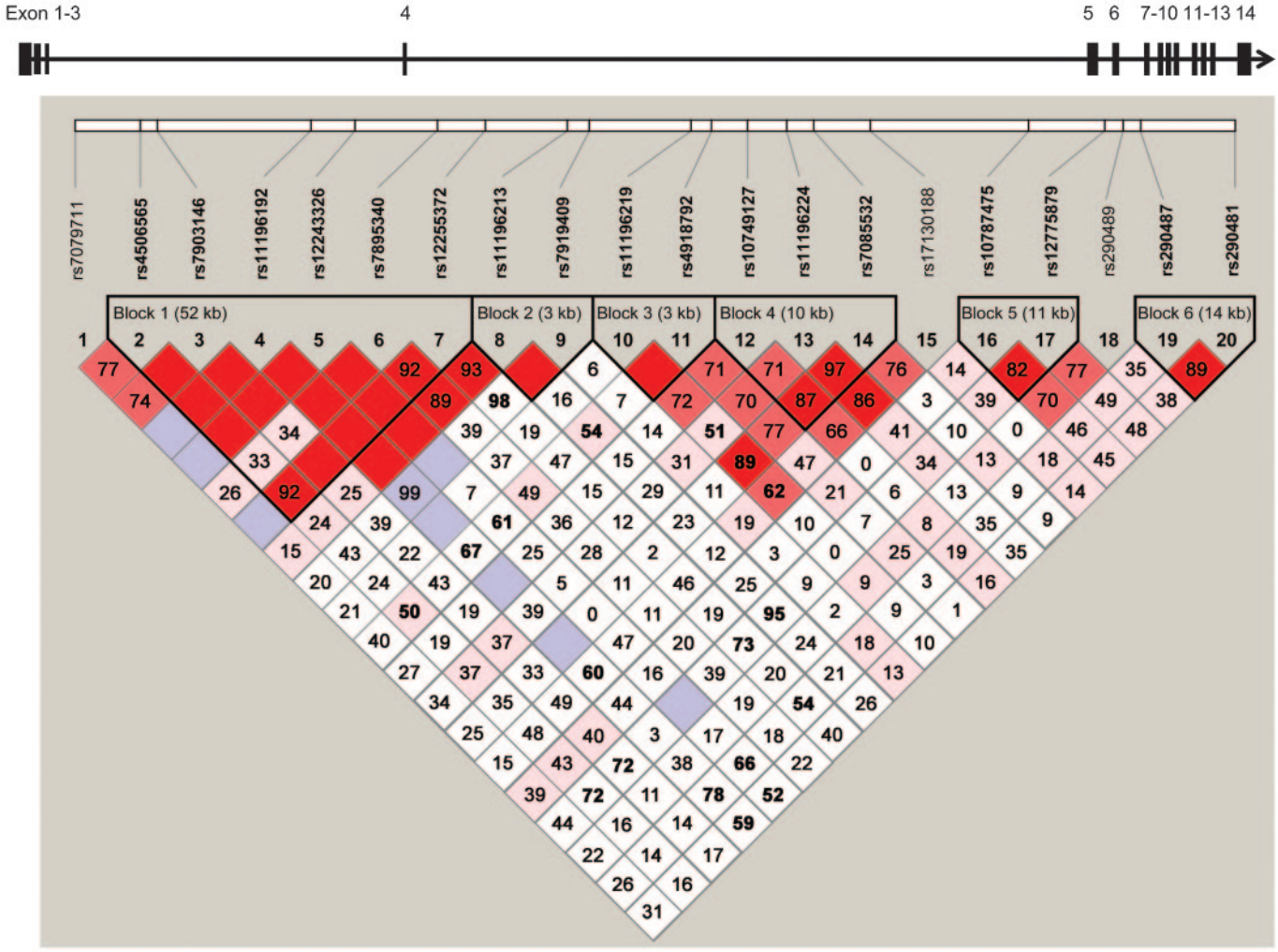

FIG. 1. Graphical representation of SNPs in relation to the exon-intron structure (*upper line*) and Haploview LD graph of *TCF7L2* gene (*lower panel*). The exon regions are shown with filled rectangles and numbered in order. Pairwise LD coefficients D′ × 100 are shown in each cell (D′ values of 1.0 are not shown). Standard color scheme of Haploview was applied for LD color display (logarithm of odds [LOD] score ≥2 and D′ = 1, shown in bright red; LOD score ≥2 and D′ <1 shown in blue; LOD score <2 and D′ = 1 shown in pink; LOD score <2 and D′ <1 shown in white).

TABLE 1
Clinical and biochemical characteristics of type 2 diabetic case subjects and normoglycemic control subjects

	Control subjects	Case subjects	*P* value
n	760	760	
Age (years)	64.52 ± 13.64	60.03 ± 11.83	<0.001
Male:female ratio	382:378	429:331	0.016
BMI (kg/m^2)	23.61 ± 3.16	24.66 ± 3.40	<0.001
Fasting glucose (mmol/l)	5.37 ± 0.73	8.65 ± 2.81	<0.001
A1C (%)	5.44 ± 0.56	7.90 ± 1.66	<0.001
Triglycerides (mmol/l)	1.27 ± 0.73	1.89 ± 1.89	<0.001
Total cholesterol (mmol/l)	5.05 ± 0.97	5.25 ± 1.31	0.0012

Data are means ± SD.

multivariate analysis by considering age, sex, and BMI as covariates using the Haplo.Stats program. The population attributable risk fraction (PAF) was estimated with the data from the control group, calculated as follows: $1 - \{1/[p^2\ \mathrm{OR}_{\mathrm{homo}} + 2p(1-p)\mathrm{OR}_{\mathrm{hetero}} + (1-p)^2]\}$, where p is the risk-allele frequency, $\mathrm{OR}_{\mathrm{homo}}$ is the odds ratio (OR) for homozygotes, and $\mathrm{OR}_{\mathrm{hetero}}$ is the OR for heterozygotes.

RESULTS

Characteristics of subjects. We genotyped a total of 1,520 unrelated subjects (including 760 type 2 diabetic case and 760 control subjects) from a Han Chinese population in Taiwan. The Han Chinese is the largest ethnic group in Taiwan, comprising ~98% of the general population. The clinical and biochemical characteristics are summarized in Table 1.

Characteristics of the SNPs and LD structures of *TCF7L2* gene. On average, 99.08% of attempted genotypes were successful (success rates ranging from 98.1 to 99.9% for each SNP) except for rs6585194, in which genotyping failed in all samples. The concordance rate of genotyping duplication was 99.93%. Graphical representation of SNPs in relation to the exon-intron structure is shown in Fig. 1. All SNPs were in introns. The genomic position, nucleic acid composition, and minor allele frequencies of the 20 genotyped SNPs are summarized in Table 2.

To determine the extent of LD in the *TCF7L2* gene, genotype data of both case and control groups were used to estimate intermarker LD. Standardized pairwise LD coefficients D′ and r^2 between markers were estimated (Table 3). Six LD blocks were identified across the gene (Fig. 1). LD block 1 spanning over a part of intron 3, the whole exon 4, and a part of the intron 4 corresponds to the exon 4 LD block described previously by Grant et al. (6), which contains the variants of the strongest associations with type 2 diabetes in the populations of European ancestry.

Association analysis of genetic variants of *TCF7L2* gene with type 2 diabetes. The minor allele frequencies of SNPs located in LD block 1 were low (0.4–2.6%, Table 2) in this Chinese population compared with populations of European ancestry. The genetic variations within this region were rare, with one single haplotype (A-C-T-T-G-G) accounting for 95% of all chromosomes (Table 4). We did not detect any association of type 2 diabetes with any SNP within this LD block, including rs7903146 and rs12255372 (Table 2). There was no association of haplotypes in LD

TABLE 2
TCF7L2 sequence variants and association with type 2 diabetes

No.	SNP name	Gene position (kb)	Gene region	Minor/major allele	MAF		OR (95% CI)	Allelic *P* (adjusted *P**)	Permuted *P*†
					Case	Control			
1	rs7079711	35.57	Intron 3	A/G	0.0266	0.0276	0.96 (0.60–1.54)	0.86 (0.92)	1.00
2	rs4506565	45.82	Intron 3	T/A	0.0227	0.0273	0.83 (0.51–1.35)	0.42 (0.70)	0.99
3	rs7903146	48.13	Intron 3	T/C	0.0234	0.0287	0.81 (0.50–1.31)	0.36 (0.70)	0.99
4	rs11196192	72.07	Intron 3	G/T	0.0040	0.0020	2.00 (0.43–12.4)	0.31 (0.41)	0.99
5	rs12243326	78.59	Intron 3	C/T	0.0050	0.0040	1.34 (0.41–4.71)	0.58 (0.76)	1.00
6	rs7895340	91.31	Intron 4	A/G	0.0260	0.0208	1.25 (0.76–2.10)	0.35 (0.27)	0.99
7	rs12255372	98.68	Intron 4	T/G	0.0060	0.0040	1.51 (0.48–5.16)	0.43 (0.53)	0.99
8	rs11196213	111.34	Intron 4	T/C	0.0246	0.0174	1.43 (0.84–2.47)	0.16 (0.17)	0.89
9	rs7919409	114.76	Intron 4	C/T	0.2540	0.2427	1.06 (0.90–1.26)	0.47 (0.71)	0.99
10	rs11196219	130.42	Intron 4	A/G	0.2750	0.2444	1.17 (0.99–1.39)	0.055 (0.16)	0.51
11	rs4918792	133.73	Intron 4	G/A	0.4926	0.4854	1.03 (0.89–1.19)	0.69 (0.72)	1.00
12	rs10749127	139.13	Intron 4	T/C	0.2977	0.2779	1.10 (0.94–1.15)	0.22 (0.59)	0.96
13	rs11196224	145.18	Intron 4	T/C	0.3081	0.3123	0.98 (0.84–1.15)	0.80 (0.32)	1.00
14	rs7085532	149.24	Intron 4	A/G	0.4321	0.4234	1.04 (0.90–1.20)	0.63 (0.98)	1.00
15	rs17130188	157.94	Intron 4	C/T	0.4721	0.4474	1.11 (0.96–1.28)	0.15 (0.10)	0.87
16	rs10787475	182.25	Intron 4	C/T	0.2895	0.2633	1.14 (0.99–1.34)	0.10 (0.08)	0.78
17	rs12775879	193.98	Intron 6	G/T	0.2430	0.2063	1.23 (1.04–1.49)	0.016 (0.023)	0.19
18	rs290489	196.86	Intron 7	A/G	0.3243	0.3114	1.06 (0.91–1.24)	0.44 (0.46)	0.99
19	rs290487	199.51	Intron 7	C/T	0.4178	0.3632	1.26 (1.09–1.47)	0.0021 (0.0035)	0.030
20	rs290481	213.61	Intron 13	G/A	0.4178	0.3773	1.19 (1.02–1.38)	0.021 (0.024)	0.27

*Adjusted for age, sex, and BMI. †Permutation 10,000 times. MAF, minor allele frequency.

TABLE 3
Standardized pairwise LD coefficients D′ (below diagonal of empty cells) and r^2 (above diagonal) of polymorphic sites at the *TCF7L2* gene locus

											r^2 →									
	SNP no.	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	SNP12	SNP13	SNP14	SNP15	SNP16	SNP17	SNP18	SNP19	SNP20
D′ ↓	SNP1	—	0.55	0.53	0	0	0.061	0	0.017	0	0.004	0.005	0.005	0.007	0.002	0.001	0.011	0.002	0.001	0.001	0.002
	SNP2	0.78	—	0.97	0.12	0.17	0.10	0.16	0.047	0.002	0.004	0.007	0.003	0.008	0.004	0.005	0.013	0.004	0	0	0
	SNP3	0.74	**1**	—	0.11	0.18	0.11	0.19	0.05	0.001	0.004	0.005	0.003	0.008	0.004	0.006	0.012	0.004	0	0	0.001
	SNP4	**1**	**1**	**1**	—	0.64	0.13	0.60	0.14	0.001	0.001	0.001	0.001	0	0.002	0.001	0.002	0	0.001	0.003	0.002
	SNP5	**1**	**1**	**1**	**1**	—	0.18	0.93	0.22	0.002	0	0.002	0	0	0	0.001	0	0.001	0	0.003	0.002
	SNP6	0.27	0.34	0.34	**1**	**1**	—	0.17	0.69	0.001	0.001	0.006	0.001	0.001	0	0	0.003	0.001	0.002	0.001	0.002
	SNP7	**1**	**0.93**	**1**	**1**	**1**	**0.93**	—	0.20	0.002	0	0.001	0	0	0	0.001	0.001	0.001	0	0.002	0.001
	SNP8	0.15	0.24	0.25	**1**	**1**	**0.90**	**0.93**	—	0.007	0	0.007	0	0.012	0.002	0	0.003	0.006	0.003	0.002	0.002
	SNP9	0.21	0.44	0.40	**0.99**	**1**	0.39	**0.99**	**1**	—	0.001	0.002	0.003	0.015	0.003	0.014	0.001	0.001	0	0.017	0.009
	SNP10	0.22	0.24	0.22	**1**	0.077	0.37	0.19	0.17		—	0.36	0.45	0.21	0.37	0.12	0.001	0	0.007	0.002	0.003
	SNP11	0.40	0.51	0.43	0.68	0.61	0.49	0.47	0.54	0.068	**1**	—	0.20	0.23	0.47	0.18	0.018	0.002	0.03	0.001	0
	SNP12	0.27	0.20	0.20	**1**	0.25	0.36	0.15	0.15	0.079	0.73	0.72	—	0.46	0.41	0.15	0	0.003	0.006	0.01	0.007
	SNP13	0.34	0.38	0.37	0.39	0.054	0.29	0.12	0.30	0.15	0.52	0.70	0.71	—	0.57	0.29	0.029	0.015	0.004	0.036	0.037
	SNP14	0.25	0.35	0.33	**1**	0	0.11	0.025	0.24	0.32	**0.89**	0.78	**0.88**	**0.97**	—	0.38	0.001	0.002	0.011	0.005	0.004
	SNP15	0.15	0.48	0.49	0.61	0.48	0.12	0.47	0.13	0.12	0.62	0.47	0.66	**0.87**	0.77	—	0.009	0.055	0	0.026	0.016
	SNP16	0.39	0.44	0.41	0.45	0.16	0.20	0.20	0.25	0.19	0.1	0.21	0.008	0.41	0.039	0.14	—	0.52	0.41	0.13	0.12
	SNP17	0.44	0.73	0.73	0.03	**1**	0.40	0.74	**0.95**	0.03	0.009	0.074	0.066	0.34	0.11	0.40	**0.82**	—	0.38	0.11	0.11
	SNP18	0.22	0.16	0.11	0.39	0.18	0.20	0.21	0.24	0.099	0.094	0.25	0.085	0.14	0.13	0.004	0.70	0.78	—	0.09	0.10
	SNP19	0.26	0.14	0.14	0.79	0.67	0.18	0.54	0.21	0.026	0.099	0.033	0.19	0.35	0.098	0.19	0.46	0.50	0.35	—	0.78
	SNP20	0.311	0.17	0.17	0.60	0.52	0.23	0.41	0.27	0.18	0.11	0.016	0.16	0.35	0.092	0.14	0.46	0.49	0.38	**0.90**	—

*Pairwise LD coefficients D′ >0.8 are shown in bold.

block 1 with type 2 diabetes (global $P = 0.47$, best haplotype-specific $P = 0.39$) (Table 4). Sliding window analysis for the haplotypes composed of adjacent SNPs (2–10 SNPs in length) also failed to reveal any association with type 2 diabetes within this region (Table 5).

However, we found that SNP rs290487 (minor allele frequency 39.03%), located 150 kb downstream of rs7903146, was significantly associated with type 2 diabetes (nominal $P = 0.0021$) (Table 2). The association remained significant after correction for multiple testing (permuted $P = 0.03$) (Table 2). The allelic OR of the rs290487 C-allele carriers was 1.26 (95% CI 1.08–1.46) (Table 2). The genotypic OR for the heterozygous genotype (CT) was 1.36 (1.08–1.71, $P = 0.0063$) and 1.51 (1.10–2.07, $P = 0.0085$) for homozygous genotype (CC), compatible with either dominant or additive genetic models ($P = 0.0017$ and 0.0059, respectively) and corresponding to a PAF of 18.7%. The SNP rs290487 was located in an LD block (LD block 6) spanning from intron 7 to intron 13 (14 kb in length) (Fig. 1). Another SNP rs290481 in this block was also associated with type 2 diabetes with marginal significance (nominal $P = 0.021$, permuted $P = 0.27$) (Table 2). The haplotypes within LD block 6 were significantly associated with type 2 diabetes (global $P = 0.012$) (Table 4). Among them, the haplotype T-A was associated with a reduced risk (OR 0.83 [95% CI 0.71–0.96], permuted $P = 0.033$) (Table 4), and the haplotype C-G was associated with an increased risk of type 2 diabetes (1.24 [1.06–1.44], permuted $P = 0.014$) (Table 4). These associations remained significant after adjustment for sex, age, and BMI (Tables 2 and 4), suggesting their effects on diabetes were not primarily mediated through adiposity. Sliding window analysis also revealed consistently significant associations of haplotypes containing rs290487 with type 2 diabetes (best global $P = 0.002$) (Table 5).

DISCUSSION

The polymorphism of *TCF7L2* (rs7903146) is the strongest single genetic variant associated with type 2 diabetes (6,23–26) and has been convincingly replicated in multiple populations (7–26). In this study, we aimed to explore the effect of the *TCF7L2* polymorphisms in a Han Chinese population. We found that the type 2 diabetes risk–conferring alleles in the exon 4 LD block described previously by Grant et al. (6), such as rs7903146 T and rs12255372 T, were rare in this Chinese population. We did not detect any association of genetic variants within this region with type 2 diabetes in this Chinese population. However, given their low allele frequencies of 2.3% and assuming the genotype relative risks of 1.45 and 2.41 in heterozygous and homozygous carriers, respectively (6), and the diabetes prevalence of 8% in the Chinese population in Taiwan, enrollment of ~1,700 case subjects would be necessary for the case-control study to have 80% power (32). Thus, a sample size at least twice larger than our study would be needed to detect the effect of these genetic variants.

Florez et al. (22) recently reported the associations of these variants with increased risk of developing diabetes in a large prospective cohort of the Diabetes Prevention Program. In subgroup analysis, there were also no associations of these variants with type 2 diabetes in Asians, which could at least attribute to a small sample size (128 participants) (22). Horikoshi et al. (33) and Hayashi et al. (34) independently reported significant associations be-

TABLE 4
Association analysis of TCF7L2 haplotypes with type 2 diabetes

Haplotypes	Frequencies		OR (95% CI)	Nominal *P* (adjusted *P**)	Permuted *P*†	Global *P* (adjusted *P**)
	Case	Control				
Block 1						0.47 (0.58)
A-C-T-T-G-G	0.947	0.959	1.03 (0.71–1.51)	0.85 (0.55)	0.99	
T-T-T-T-G-G	0.014	0.019	0.68 (0.37–1.26)	0.20 (0.62)	0.39	
A-C-T-T-A-G	0.017	0.012	1.26 (0.67–2.40)	0.39 (0.20)	0.70	
Block 2						0.26 (0.36)
C-T	0.726	0.745	0.91 (0 0.77–1.07)	0.25 (0.43)	0.46	
C-C	0.256	0.244	1.07 (0.90–1.26)	0.47 (0.70)	0.73	
T-T	0.024	0.017	1.44 (0.84–2.49)	0.17 (0.17)	0.35	
Block 3						0.23 (0.32)
G-A	0.506	0518	0.97 (0 0.84–1.12)	0.66 (0.73)	0.90	
A-G	0.274	0246	1.17 (0 0.99–1.38)	0.061 (0.17)	0.14	
G-G	0.219	0.242	0.88 (0 0.74–1.05)	0.15 (0.31)	0.33	
Block 4						0.12 (0.11)
C-C-G	0.543	0.554	0.97 (0 0.84–1.12)	0.68 (0.87)	0.99	
T-T-A	0.230	0.233	0.99 (0 0.84–1.18)	0.93 (0.44)	1.00	
C-C-A	0.082	0.089	0.92 (0.71–1.20)	0.52 (0.62)	0.98	
C-T-A	0.075	0.077	0.98 (0.74–1.29)	0.88 (0.82)	1.00	
T-C-A	0.047	0.029	1.62 (1.09–2.44)	0.011 (0.011)	0.04	
T-C-G	0.021	0.019	1.11 (0.64–1.93)	0.70 (0.51)	0.99	
Block 5						0.16 (0.20)
T-T	0.679	0.714	0.87 (0.74–1.01)	0.072 (0.05)	0.20	
C-G	0.212	0.181	1.22 (1.01–1.46)	0.034 (0.06)	0.10	
C-T	0078	0.083	0.94 (0 0.72–1.24)	0.65 (0.97)	0.95	
T-G	0.031	0.027	1.18 (0.75–1.86)	0.50 (0.39)	0.87	
Block 6						0.012 (0.025)
T-A	0.555	0.601	0.83 (0 0.71–0.96)	0.011 (0.0092)	0.033	
C-G	0.390	0.340	1.24 (1.06–1.44)	0.0046 (0.012)	0.014	
T-G	0.027	0.036	0.74 (0.48–1.14)	0.15 (0.42)	0.42	
C-A	0.027	0.022	1.25 (0.77–2.06)	0.34 (0.21)	0.72	

*Adjusted for age, sex, and BMI. †Permutation 10,000 times.

tween rs7903146 and type 2 diabetes in the Japanese population. However, the minor allele frequency of rs7903146 was low (3–5%) in their study. The PAF for rs7903146 is only ~3% in the Japanese population, a much lower value than that in the populations of European ancestry (6,33,34). These data indicated that the genetic variants within the exon 4 LD block were rare among the Japanese and Chinese population, so that large sample sizes would be needed to detect the effects.

Interestingly, we identified a novel association of SNP

TABLE 5
Sliding window association analysis of global *P* value* for haplotypes composed of adjacent 2–10 SNPs

No.	SNP	2 SNPs	3 SNPs	4 SNPs	5 SNPs	6 SNPs	7 SNPs	8 SNPs	9 SNPs	10 SNPs
1	rs7079711	0.20	0.26	0.25	0.33	0.48	0.47	0.40	0.47	0.11
2	rs4506565	0.69	0.28	0.40	0.46	0.47	0.33	0.40	0.078	0.16
3	rs7903146	0.28	0.39	0.47	0.38	0.25	0.33	0.087	0.16	0.22
4	rs11196192	0.50	0.66	0.54	0.44	0.55	0.099	0.10	0.18	0.098
5	rs12243326	0.67	0.54	0.44	0.54	0.12	0.11	0.18	0.088	0.049
6	rs7895340	0.53	0.45	0.54	0.12	0.068	0.17	0.087	0.050	0.027
7	rs12255372	0.31	0.39	0.076	0.068	0.16	0083	0.049	0.027	0.11
8	rs11196213	0.26	0.083	0.11	0.16	0.070	0.034	0.026	0.089	0.023
9	rs7919409	0.057	0.10	0.20	0.07	0.30	0.038	0.15	0.028	0.057
10	rs11196219	0.12	0.48	0.33	0.30	0.096	0.11	0.062	0.075	0.003
11	rs4918792	0.78	0.44	0.35	0.096	0.46	0.13	0.098	0.017	0.017
12	rs10749127	0.14	0.12	0.25	0.46	0.22	0.15	0.026	0.030	
13	rs11196224	0.40	0.31	0.51	0.23	0.086	0.017	0.034		
14	rs7085532	0.13	0.12	0.20	0.086	0.004	0.015			
15	rs17130188	0.077	0.15	0.10	0.004	0.045				
16	rs10787475	0.16	0.18	0.017	0.045					
17	rs12775879	0.085	0.002	0.024						
18	rs290489	0.022	0.15							
19	rs290487	0.012								
20	rs290481									

*Permutation 10,000 times.

rs290487 C located in an LD bock spanning over intron 7 to intron 13 with type 2 diabetes. In contrast to rs7903146 T, rs290487 C was common (allele frequency 39.03%) in the Chinese population. The risk allele and genotypes were associated with a significantly increased risk of type 2 diabetes, and the corresponding PAF was also substantial (18.7%) in the Chinese population, comparable with that of rs7903146 T in the populations of European origin (6). These data suggested that genetic variants of the *TCF7L2* gene were also major determinants of type 2 diabetes in the Chinese population. Moreover, our results along with the published reports from other ethnic groups demonstrate that different variations of the *TCF7L2* gene confer the risk of type 2 diabetes in different populations. These findings provided novel evidence supporting a role of the *TCF7L2* locus in the pathogenesis of type 2 diabetes in a Chinese population.

The underlying mechanisms by which a genetic variation within the intron of the *TCF7L2* gene confers susceptibility of type 2 diabetes remain to be elucidated. The LD block harboring rs290487 lies very close (<2 kb) to the 3′ end of *TCF7L2* gene. The human *TCF7L2* gene displays a high degree of alternative splicing near the 3′ end of the gene (35,36). The 3′ end of the *TCF7L2* gene encodes the binding domains for COOH-terminal binding protein, a protein implicated in the repression of the *TCF7L2* transcriptional activity. The alternative splicing leads to the synthesis of a number of isoforms with different COOH-terminal ends and binding domain activities, which were associated with altered downstream signaling activities (35,36). Furthermore, many alternative splice sites were observed either experimentally or in silico in this region (35). It is intriguing to speculate that genetic variations in this region would affect the action of *TCF7L2* through the regulation of alternative splicing. Based on multispecies sequence alignment analysis using a phylogenic hidden Markov model implemented in the PhastCons program (http://genome.ucsc.edu/cgi-bin/hgGateway) (37), there is a high degree of evolutional conservation in the 3′ end of human *TCF7L2* gene, which we identified for association of diabetes. Functional studies will be required to determine the relationship between sequence variations, gene expression, protein product, and function in human tissues. LD block 6 contains two nonsynonymous exonic SNPs and 37 intronic SNPs. Therefore, further extensive search within this region may be needed to identify the true risk-conferring site as well as to confirm the association.

In summary, we identified a novel association of the genetic variant SNP rs290487 in the *TCF7L2* gene with type 2 diabetes in a Chinese population, in contrast to the previous reports of the SNP rs7903146 in other ethnic groups. The variant was located in an LD block close to the 3′ end of the *TCF7L2* gene, a region displaying a high degree of alternative splicing. The effect of this variant was significant, with a substantial PAF in the Chinese population. We did not detect any association of the previously established risk allele (rs7904136 T and rs12255372 T) with type 2 diabetes, probably owing to their low frequency in the Chinese population. These data suggested that genetic polymorphisms of the *TCF7L2* gene were associated with type 2 diabetes and are major genetic contributors of the risk of type 2 diabetes a Chinese population.

ACKNOWLEDGMENTS

This research project was supported by grants from the National Science Council (NSC-93-B-002-005 and NSC-94-3112-B-002-019) of Taiwan.

We thank all of the patients for their participation in this study.

REFERENCES

1. Barroso I: Genetics of type 2 diabetes. *Diabet Med* 22:517–535, 2005
2. Hattersley AT, McCarthy MI: What makes a good genetic association study? *Lancet* 366:1315–1323, 2005
3. Altshuler D, Hirschhorn JN, Klannemark M: The common PPARg Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76–80, 2000
4. Gloyn AL, Weedon MN, Owen KR: Large-scale association studies of variants in genes encoding the pancreatic β-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* 52:568–572, 2003
5. Weedon MN, Schwarts PE, Horikawa Y: Meta-analyses and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. *Am J Hum Genet* 73:1208–1212, 2003
6. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Hegason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38:320–323, 2006
7. Lehman DM, Hunt KJ, Leach RJ, Hamlington J, Arya R, Abboud HE, Duggirala R, Blangero J, Goring HH, Stern MP: Haplotypes of transcription factor 7-like 2 (TCF7L2) gene and its upstream region are associated with type 2 diabetes and age of onset in Mexican Americans. *Diabetes* 56:389–393, 2007
8. Mayans S, Lackovic K, Lindgren P, Ruikka K, Agren A, Eliasson H, Holmberg D: TCF7L2 polymorphisms are associated with type 2 diabetes in northern Sweden. *Eur J Hum Genet* 15:342–346, 2007
9. Marzi C, Huth C, Kolz M, Grallert H, Meisinger C, Wichmann HE, Rathmann W, Herder C, Illig T: Variants of the transcription factor 7-like 2 gene (TCF7L2) are strongly associated with type 2 diabetes but not with the metabolic syndrome in the MONICA/KORA Surveys. *Horm Metab Res* 39:46–52, 2007
10. Munoz J, Lok KH, Gower BA, Fernandez JR, Hunter GR, Lara-Castro C, De Luca M, Garvey WT: Polymorphism in the transcription factor 7-like 2 (TCF7L2) gene is associated with reduced insulin secretion in nondiabetic women. *Diabetes* 55:3630–3634, 2006
11. Chandak GR, Janipalli CS, Bhaskar S, Kulkarni SR, Mohankrishna P, Hattersley AT, Frayling TM, Yajnik CS: Common variants in the TCF7L2 gene are strongly associated with type 2 diabetes mellitus in the Indian population. *Diabetologia* 50:63–67, 2007
12. Cauchi S, Meyre D, Choquet H, Dina C, Born C, Marre M, Balkau B, Froguel P, the DESIR Study Group: TCF7L2 variation predicts hyperglycemia incidence in a French general population: the data from an epidemiological study on the Insulin Resistance Syndrome (DESIR) study. *Diabetes* 55:3189–3192, 2006
13. Humphries SE, Gable D, Cooper JA, Ireland H, Stephens JW, Hurel SJ, Li KW, Palmen J, Miller MA, Cappuccio FP, Elkeles R, Godsland I, Miller GJ, Talmud PJ: Common variants in the TCF7L2 gene and predisposition to type 2 diabetes in UK European Whites, Indian Asians and Afro-Caribbean men and women. *J Mol Med* 84 (Suppl. 12):1–10, 2006
14. van Vliet-Ostaptchouk JV, Shiri-Sverdlov R, Zhernakova A, Strengman E, van Haeften TW, Hofker MH, Wijmenga C: Association of variants of transcription factor 7-like 2 (TCF7L2) with susceptibility to type 2 diabetes in the Dutch Breda cohort. *Diabetologia* 50:59–62, 2007
15. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, Rayner NW, Shields B, Owen KR, Hattersley AT, Frayling TM: Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med* 3:e374, 2006
16. Cauchi S, Meyre D, Dina C, Choquet H, Samson C, Gallina S, Balkau B, Charpentier G, Pattou F, Stetsyuk V, Scharfmann R, Staels B, Fruhbeck G, Froguel P: Transcription factor TCF7L2 genetic study in the French population: expression in human β-cells and adipose tissue and strong association with type 2 diabetes. *Diabetes* 55:2903–2908, 2006
17. Saxena R, Gianniny L, Burtt NP, Lyssenko V, Giuducci C, Sjogren M, Florez JC, Almgren P, Isomaa B, Orho-Melander M, Lindblad U, Daly MJ, Tuomi T, Hirschhorn JN, Ardlie KG, Groop LC, Altshuler D: Common single

nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals. *Diabetes* 55:2890–2895, 2006
18. Damcott CM, Pollin TI, Reinhart LJ, Ott SH, Shen H, Silver KD, Mitchell BD, Shuldiner AR: Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance. *Diabetes* 55:2654–2659, 2006
19. Scott LJ, Bonnycastle LL, Willer CJ, Sprau AG, Jackson AU, Narisu N, Duren WL, Chines PS, Stringham HM, Erdos MR, Valle TT, Tuomilehto J, Bergman RN, Mohlke KL, Collins FS, Boehnke M: Association of transcription factor 7-like 2 (TCF7L2) variants with type 2 diabetes in a Finnish sample. *Diabetes* 55:2649–2653, 2006
20. Zhang C, Qi L, Hunter DJ, Meigs JB, Manson JE, van Dam RM, Hu FB: Variant of transcription factor 7-like 2 (TCF7L2) gene and the risk of type 2 diabetes in large cohorts of U.S. women and men. *Diabetes* 55:2645–2648, 2006
21. Groves CJ, Zeggini E, Minton J, Frayling TM, Weedon MN, Rayner NW, Hitman GA, Walker M, Wiltshire S, Hattersley AT, McCarthy MI: Association analysis of 6,736 U.K. subjects provides replication and confirms TCF7L2 as a type 2 diabetes susceptibility gene with a substantial effect on individual risk. *Diabetes* 55:2640–2644, 2006
22. Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PI, Shuldiner AR, Knowler WC, Nathan DM, Altshuler D: Diabetes Prevention Program Research Group: TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med* 355:241–250, 2006
23. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881–885, 2007
24. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University and Norvatis Institutes for BioMedical Research: Genome-wise association analysis identified loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336, 2007
25. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS, Wellcome Trust Case Control Consortium (WTCCC), McCarthy MI, Hattersley AT: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1339–1341, 2007
26. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1336–1339, 2007
27. Helgason A, Palsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, Benediktsson R, Hinney A, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Schafer H, Faruque M, Doumatey A, Zhou J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Sigurdsson G, Hebebrand J, Pedersen O, Thorsteinsdottir U, Gulcher JR, Kong A, Rotimi C, Stefansson K: Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 39:218–225, 2007
28. American Diabetes Association: Standards of medical care in diabetes. *Diabetes Care* 27 (Suppl. 1):S15–S35, 2004
29. Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265, 2005
30. Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R, Modica F, Pohl M, Varde S, Zhao R, Zhao X, Boyce-Jacino MT: SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques* 32 (Suppl.):70–77, 2002
31. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 70:425–434, 2002
32. Dupont WD, Plummer WD: Power and sample size calculations for studies involving linear regression. *Controlled Clin Trials* 19:589–601, 1998
33. Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T: A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* 50:747–751, 2007
34. Hayashi T, Iwamoto Y, Kaku K, Hirose H, Maeda S: Replication study for the association of *TCF7L2* with susceptibility to type 2 diabetes in a Japanese population. *Diabetologia* 50:980–984, 2007
35. Duval A, Rolland S, Tubacher E, Bui H, Thomas G, Hamelin R: The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. *Cancer Res* 60:3872–3879, 2000
36. Shiina H, Igawa M, Breault J, Ribeiro-Filho L, Pookot D, Urakami S, Terashima M, Deguchi M, Yamanaka M, Shirai M, Kaneuchi M, Kane CJ, Dahiya R: The human T-cell factor-4 gene splicing isoforms, Wnt signal pathway, and apoptosis in renal cell carcinoma. *Clin Cancer Res* 9:2121–2131, 2003
37. Siepel A, Haussler D: Combining phylogenetic and hidden Markov models in biosequence analysis. *J Comput Biol* 11:413–428, 2004