

# A Genome-Wide Association Study Confirms Previously Reported Loci for Type 2 Diabetes in Han Chinese

Bin Cui<sup>1,2,3</sup>, Xiaolin Zhu<sup>1,2,3</sup>, Min Xu<sup>1,2,3</sup>, Ting Guo<sup>2</sup>, Dalong Zhu<sup>3</sup>, Gang Chen<sup>4</sup>, Xuejun Li<sup>5</sup>, Lingyan Xu<sup>2</sup>, Yufang Bi<sup>2</sup>, Yuhong Chen<sup>2</sup>, Yu Xu<sup>2</sup>, Xiaoying Li<sup>1,2</sup>, Weiqing Wang<sup>2</sup>, Haifeng Wang<sup>6</sup>, Wei Huang<sup>6\*</sup>, Guang Ning<sup>1,2\*</sup>

**1** Key Laboratory for Endocrine and Metabolic Diseases of Ministry of Health, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, **2** Department of Endocrine and Metabolic Diseases, Shanghai Clinical Center for Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, **3** Department of Endocrine Diseases, Nanjing Drum Tower Hospital, Affiliated Hospital of Nanjing University Medical School, Nanjing, China, **4** Department of Endocrine Diseases, Fujian Provincial Hospital, Fuzhou, China, **5** Department of Endocrine Diseases, First Affiliated Hospital of Xiamen University, Xiamen, China, **6** Department of Genetics, Chinese National Human Genome Center, Shanghai, China

## Abstract

**Background:** Genome-wide association study (GWAS) has identified more than 30 loci associated with type 2 diabetes (T2D) in Caucasians. However, genomic understanding of T2D in Asians, especially Han Chinese, is still limited.

**Methods and Principal Findings:** A two-stage GWAS was performed in Han Chinese from Mainland China. The discovery stage included 793 T2D cases and 806 healthy controls genotyped using Illumina Human 660- and 610-Quad BeadChips; and the replication stage included two independent case-control populations (a total of 4445 T2D cases and 4458 controls) genotyped using TaqMan assay. We validated the associations of *KCNQ1* (rs163182,  $p = 2.085 \times 10^{-17}$ , OR 1.28) and *C2CD4A/B* (rs1370176,  $p = 3.677 \times 10^{-4}$ , OR 1.124; rs1436953,  $p = 7.753 \times 10^{-6}$ , OR 1.141; rs7172432,  $p = 4.001 \times 10^{-5}$ , OR 1.134) in Han Chinese.

**Conclusions and Significance:** Our study represents the first GWAS of T2D with both discovery and replication sample sets recruited from Han Chinese men and women residing in Mainland China. We confirmed the associations of *KCNQ1* and *C2CD4A/B* with T2D, with the latter for the first time being examined in Han Chinese. Arguably, eight more independent loci were replicated in our GWAS.

**Citation:** Cui B, Zhu X, Xu M, Guo T, Zhu D, et al. (2011) A Genome-Wide Association Study Confirms Previously Reported Loci for Type 2 Diabetes in Han Chinese. *PLoS ONE* 6(7): e22353. doi:10.1371/journal.pone.0022353

**Editor:** Kerby Shedden, University of Michigan, United States of America

**Received:** April 6, 2011; **Accepted:** June 20, 2011; **Published:** July 22, 2011

**Copyright:** © 2011 Cui et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by the grants from Key Laboratory for Endocrine and Metabolic Diseases of Ministry of Chinese Health (No. 1994DP131044), State Key Laboratory of Medical Genomics, China, National Key New Drug Creation and Manufacturing Program (2008ZX09312/019) and the Sector Funds of Ministry of Health (201002002). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: gning@sibs.ac.cn (GN); huangwei@chgc.sh.cn (WH)

☞ These authors contributed equally to this work.

## Introduction

Type 2 diabetes (T2D) is a complex disease hallmarked by insulin resistance and pancreatic beta-cell dysfunction [1]. T2D is becoming a major concern of global public health due to its escalating prevalence throughout the world [2]. In China, 9.7% and 15.5% of the entire population suffer from T2D and prediabetes, respectively [3]. Although overfeeding and sedentary lifestyle are claimed as the main contributors to its increasing incidence, genetic factors play a significant role in the etiology and pathogenesis of T2D, in many cases via interaction with environmental counterparts [4].

Genetic analysis of T2D and related diseases (such as obesity and monogenic diabetes) and traits (such as fasting plasma glucose levels) has greatly improved our understanding of glucose homeostasis and energy balance in both physiological and pathological conditions, some of which has brought novel preventive and therapeutic options [4]. Before the year 2007,

studies based on linkage analysis and candidate gene strategy identified only a few genetic loci of T2D [5–7]. More recently, several genome-wide association studies (GWASs) have been completed in independent population samples derived from Caucasians and Japanese, and identified a host of novel susceptibility loci of T2D [8–19]. With a more recent large-scale meta-analysis taking into account [20], these studies in total have discovered at least 38 independent susceptibility loci of T2D, many of which have been replicated in populations of different ancestries, including Han Chinese [21,22].

Extending GWAS to populations of diverse descents is valuable, because different frequencies of genetic variants and patterns of linkage disequilibrium (LD) due to population backgrounds may strongly affect the power and potential of GWAS to discover and/or refine certain genetic loci associated with disease [23]. For example, association of *KCNQ1* with T2D was first independently identified by two GWASs in Japanese [14,15], although for both studies the sample size of the discovery

stage was small and the genomic coverage of SNPs insufficiently dense [24]. However, to date, original data of GWAS of T2D in Han Chinese is still limited [25,26]. One study was conducted among Han Chinese in Taiwan that identified two additional novel loci in the protein tyrosine phosphatase receptor type D (*PTPRD*) and serine racemase (*SRR*) genes [25]. The other was performed in women from Shanghai Women's Health Study (SWHS) and Shanghai Breast Cancer Study (SBCS) [26]. Here, we reported a 2-stage study, comprising one discovery and one replication stage, in both Han Chinese men and women residing in Mainland China.

## Results

In the discovery stage, 998 T2D patients of Han Chinese descent recruited in Shanghai were genotyped using Illumina Human 660-Quad BeadChips. Our control population was derived from a glucose survey in a community in Shanghai and was genotyped using Illumina Human 660- and 610-Quad BeadChips. After stringent quality control, we obtained 561,694 SNPs that were common to both genotyping platforms with both average call rates of >99%. To ensure that case and control groups were genetically matched, in addition to close examination of their geographic origins, MDS was used to exclude population outliers, the result of which was further confirmed by PCA, which showed minimal evidence for population stratification [Figure S1, Figure S2]. 474,515 SNPs in 793 cases and 806 controls entered final statistical analysis using the Cochran-Armitage trend test to examine the genotype-phenotype association under an additive model. After genomic control (GC) with an inflation factor  $\lambda$  of 1.08 [Figure S3], the association results did not change significantly.

We selected the top 30 significantly associated SNPs representing 15 genomic regions in the discovery stage (arbitrarily defined as trend- $P \leq 10^{-4}$ ) for genotyping in an independent case-control population from Shanghai ( $n = 2620$ ) (1058 cases and 1562 controls, part of Replication 1) as a fast-track replication analysis [Table S1, Figure S4]. Among the 30 SNPs, 3 representing 2 genomic loci were replicated with the same direction of association with the discovery stage at the significance level of  $P < 0.10$  (Figure 1). Two of them, namely rs163182 and rs163184 ( $P = 2.348 \times 10^{-7}$ , OR = 1.37, and  $P = 0.04332$ , OR = 1.121, respectively), were located in the same locus in gene *KCNQJ* as previously reported. The remaining SNP, rs3773159 ( $P = 0.09591$ , OR = 1.152), is likely to represent a novel susceptibility locus for T2D in the population.

We genotyped the rest of replication 1 (1602 cases and 506 controls residing in Shanghai and Jiangsu Province) and pooled analysis (Replication 1: 2660 cases and 2068 controls) showed a nominally significant association of rs3773159 with T2D ( $P = 0.0418$ , OR = 1.137, Table 1). Further genotyping and analysis in an independent sample comprising 1785 cases and 2390 controls (Replication 2: Southern Han) identified a consistent association though without reaching significance level. Therefore, there was not enough evidence to establish the association of rs3773159 with T2D in out populations.

We checked a total of 37 genomic loci previously reported to be associated with T2D [20] in our discovery stage data, among which 42 SNPs in 26 loci were successfully genotyped. 13 SNPs representing 9 loci were significantly associated with T2D in our GWAS at a significant level of  $P < 0.05$  (Table S2): rs2793831 (proxy for rs10923931, *NOTCH2*), rs2943641 (*IRS1*), rs2120825 (proxy for rs1801282, *PPARG*), rs734312 (*WFS1*), rs896854 (*TP53INPI*), rs10906115 (*CDC123/CAMK1D*), rs7901695 (*TCF7L2*), rs7903146

(*TCF7L2*), rs1552224 (*CENTD2*), rs1370176 (*C2CD4A/B*), rs1436953 (*C2CD4A/B*), rs7172432 (*C2CD4A/B*), and rs1436955 (*C2CD4B*). If we adopt a less stringent threshold of  $P < 0.10$ , 4 more SNPs representing 3 genomic regions showed significant association with T2D in the discovery stage (Table S2): rs1111875 (*HHEX*), rs7923837 (*HHEX*), rs8050136 (*FTO*), and rs780094 (*GCKR*).

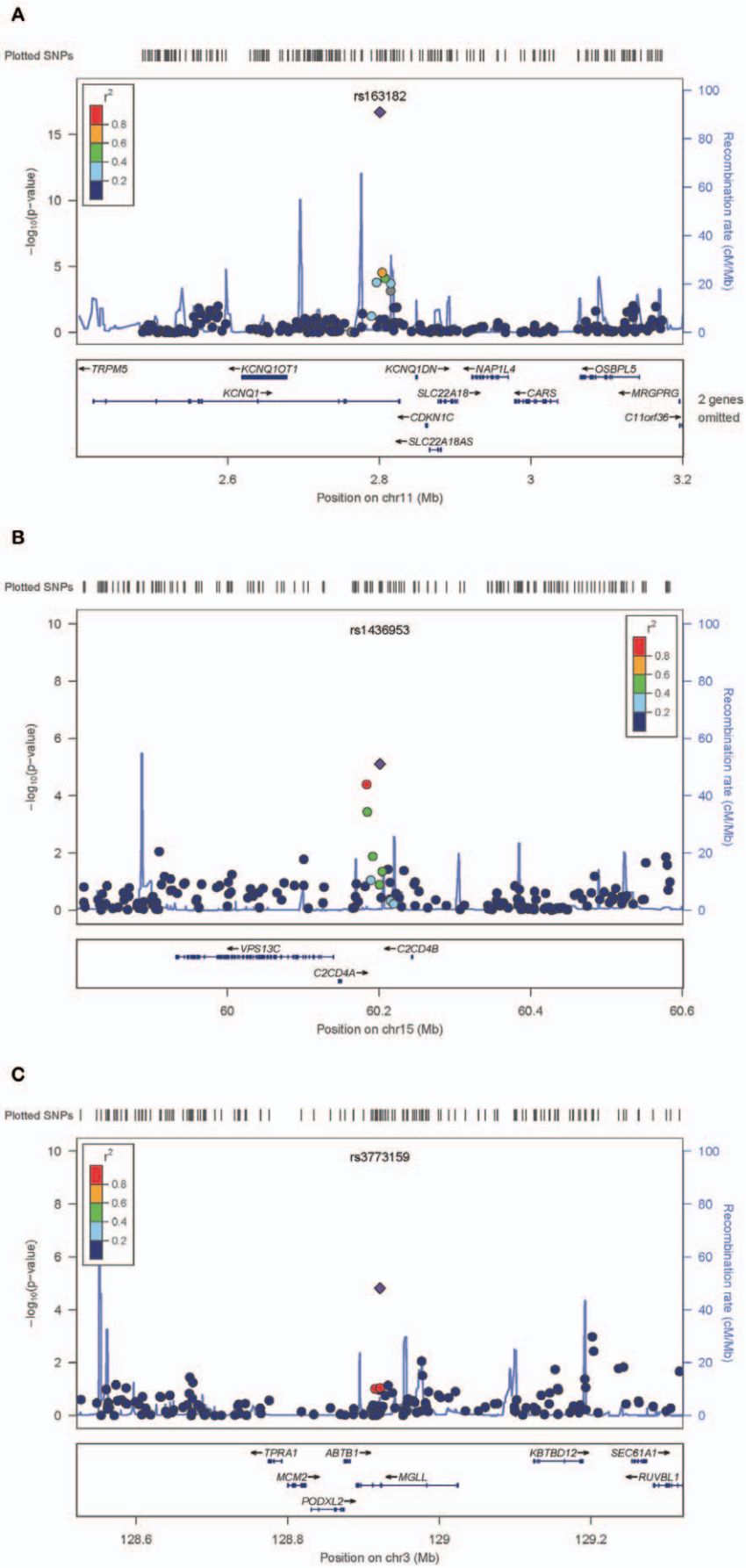
We noticed that the three SNPs – rs1370176, rs1436953, and rs7172432 in *C2CD4A/B* on chromosome 15 – were exactly the lead SNPs in a three-stage GWAS recently reported involving a total of ~19000 Japanese [27]. Further genotyping in the fast-track replication sample showed that the three SNPs tended to be consistently associated with T2D ( $P = 0.1134$ , 0.03651, and 0.01884, respectively). Genotyping in the remaining 1602 cases and 506 controls of Replication 1 and pooled analysis (Replication 1) expectedly rendered these associations significant (rs1370176,  $P = 0.03597$ , OR = 1.115; rs1436953,  $P = 9.671 \times 10^{-5}$ , OR = 1.187; rs7172432,  $P = 4.529 \times 10^{-4}$ , OR = 1.187, respectively; Table 1). Further analysis in the Southern Han population (Replication 2) showed that the associations were in the same direction as those in Central Han population but did not reach significance level (Table 1, Figure 1).

## Discussion

Our GWAS did not provide sufficient evidence for potentially novel genetic variation associated with T2D in Han Chinese. However, our study validated 10 previously reported loci associated with T2D, including *KCNQJ* and *C2CD4A/B*, and several of which were very recently discovered by large-scale multistage study and meta-analyses.

In Han Chinese population, our study validated previously known susceptibility loci of T2D, among which SNPs in *KCNQJ* were consistently and most strongly associated with T2D (rs163182,  $P_{combined} = 2.085 \times 10^{-17}$ , OR = 1.280). rs163184, the SNP most significantly associated with T2D in *KCNQJ* in our discovery stage, is also the lead SNP in two other independent studies [14,20]. One additional SNP in the same locus, namely rs2237892, was reported to be significantly associated with T2D in two independent studies in Japanese [14,15], and replicated in our previous study [22]. Collectively, these data confirmed that variants in *KCNQJ* were associated with T2D in different populations, and our GWAS worked in identifying such variants.

Aside from variants in *KCNQJ*, our study rediscovered several SNPs in additional susceptibility loci of T2D originally identified in populations of different ancestries, and SNPs in three such loci came to our notice. The first is *C2CD4A/B* in 15q21.3, which was quite recently reported by a GWAS of T2D involving about 8000 Japanese in their discovery stage [27]. The most significant SNP reported, rs7172432, was also most strongly associated with T2D at the locus in the Han Chinese population ( $P_{GWAS} = 3.264 \times 10^{-4}$ ), but did not reach the cutoff  $P = 10^{-4}$ , mainly due to our limited power because of small sample size. Two more SNPs reported, namely rs1370176 and rs1436953, were in the same locus and associated with T2D in our GWAS. We noticed that another recent work reported an association of rs1436955 with T2D in Han Chinese, which was indeed in the same locus; albeit *in silico* replication strategy might have prevented their further analysis [26]. Although not replicated in Southern Han population possibly due to population reasons, the 3 SNPs were significantly associated with T2D in Central Han population, with risk alleles having slightly stronger effect sizes (OR = 1.115–1.187) than those in Japanese population as previously reported [27]. These results present direct evidence that genetic variants in *C2CD4A/B* locus



**Figure 1. Association signals at three chromosome regions.** Regional plots displays association results for SNPs as a function of genomic distance (chromosomal position from National Center for Biotechnology Information build hg18) for *KCNQ1* (Panel A), *C2CD4A/B* (Panel B), and *MGLL* (Panel C). Top line represents genomic coverage at each locus; with each vertical tick representing genotyped SNPs. Purple diamonds indicate SNP at each locus with the strongest association. Each circle represents a SNP, with the color of the circle indicating the correlation between that SNP and the best SNP at the locus (purple diamond). Light blue lines indicate estimated recombination hot spots in HapMap. Bottom panel shows genes at each locus as annotated in the UCSC Genome Browser Annotation Database.  
doi:10.1371/journal.pone.0022353.g001

are associated with T2D in Central Han Chinese population residing in Mainland China. Further studies are required to investigate whether such associations exist in Southern and Northern Han Chinese populations and identify the causal variants.

The other two loci are *TP53LNPI* (rs896854,  $P_{GWAS} = 0.002212$ ) and *CENTD2* (rs1552224,  $P_{GWAS} = 0.04226$ ); both of them were discovered in a large-scale meta-analysis comprising more than one hundred thousand individuals of European descent (DIAGRAM+). Our GWAS was likely to validate these two newly identified loci in the Han Chinese population, but further replication is required to examine their effects. Moreover, in light of the small sample size of our GWAS stage, these results support

the notion that GWAS in Han Chinese has potential to identify novel risk loci of T2D.

Our study represents the first GWAS with both discovery and replication sample sets recruited both Han Chinese men and women residing in Mainland China. Han Chinese is geographically and genetically heterogeneous and has subpopulation structures, which may have considerable effect on design and interpretation of GWAS [28]. Allele effects in Southern Han Chinese were consistently weaker than those in Central Han Chinese (Table 1); we consider this as a result of different population backgrounds. Because T2D-associated variants show much weaker effects than alleles associated with other diseases and traits (e.g., autoimmune disease and malignancies), a much larger

**Table 1. SNPs associated with T2D in the Han Chinese population.**

	risk allele	sample size (T2D/NC)	genotyping rate	RAF in T2D	RAF in NC	OR 95% confidence interval	trend-P unadjusted
<b><i>KCNQ1</i></b>							
rs163182							
GWAS	C	793/806		0.402	0.332	1.353	$4.51 \times 10^{-5}$
Central Han	C	2660/2068	98.35%	0.387	0.326	1.305 (1.198–1.423)	$3.543 \times 10^{-9}$
Southern Han	C	1785/2390	99.04%	0.411	0.362	1.227 (1.122–1.342)	$1.205 \times 10^{-5}$
Combined	C	5238/5264		0.397	0.343	1.280	$2.085 \times 10^{-17}$
<b><i>C2CD4A/B</i></b>							
rs1370176							
GWAS	C	793/806		0.731	0.687	1.238	0.005546
Central Han	C	2158/1562	99.14%	0.720	0.698	1.115 (1.007–1.234)	0.03597
Southern Han	C	1785/2390	98.37%	0.739	0.722	1.088 (0.986–1.201)	0.09241
Combined	C	4736/4758		0.729	0.708	1.124	$3.677 \times 10^{-4}$
rs1436953							
GWAS	G	793/806		0.668	0.617	1.246	0.002698
Central Han	G	2660/2068	98.73%	0.641	0.601	1.187 (1.091–1.292)	$9.671 \times 10^{-5}$
Southern Han	G	1785/2390	99.14%	0.698	0.688	1.045 (0.950–1.148)	0.3535
Combined	G	5238/5264		0.664	0.643	1.141	$7.753 \times 10^{-6}$
rs7172432							
GWAS	A	793/806		0.644	0.582	1.302	$3.264 \times 10^{-4}$
Central Han	A	2158/1562	98.76%	0.630	0.589	1.187 (1.080–1.306)	$4.529 \times 10^{-4}$
Southern Han	A	1785/2390	98.42%	0.666	0.661	1.025 (0.993–1.058)	0.6154
Combined	A	4736/4758		0.646	0.624	1.134	$4.001 \times 10^{-5}$
<b><i>MGLL</i></b>							
rs3773159							
GWAS	T	793/806		0.183	0.107	1.879	$2.64 \times 10^{-7}$
Central Han	T	2660/2068	97.95%	0.134	0.120	1.137 (1.004–1.287)	0.0418
Southern Han	T	1785/2390	99.33%	0.141	0.134	1.061 (0.9354–1.204)	0.3507
Combined	T	5238/5264		0.144	0.124	1.196	$1.552 \times 10^{-5}$

RAF(T2D) and RAF(NC), risk allele frequency in T2D cases and controls, respectively. OR, odds ratio for risk allele.  
doi:10.1371/journal.pone.0022353.t001

sample consisting of homogenous individuals is required for genuine associations to achieve genome-wide significance. Though our initial sample size is small, which might have prevented us to discover more potentially novel associations, our GWAS has successfully replicated 10 previously reported T2D susceptibility loci, several of which are very recently discovered by large-scale studies and meta-analyses in Caucasians and Japanese. This fact lends convincing evidence of the soundness of our study, and highlights the potential of discovering novel genetic variations associated with T2D by extending GWAS to diverse populations.

In conclusion, our genome-wide association study confirmed several T2D susceptibility loci previously identified in Caucasians and Japanese, among which variants in *KCNQ1* showed the strongest association, and variants in *C2CD4A/B* were first replicated in Han Chinese residing in Mainland China.

## Methods

### Ethics statement

This study was approved by the Institutional Review Board of the Ruijin Hospital, Shanghai Jiao Tong University School of Medicine and was in accordance with the principle of the Helsinki Declaration II. The written informed consent was obtained from each participant.

### Study populations

The GWAS genotyped 998 T2D patients recruited from outpatient departments of Ruijin Hospital, and 1002 healthy controls obtained from one glucose survey in Youyi community, Baoshan district, Shanghai [29]. The replication analysis included two independent populations (Replication 1: 2660 T2D cases and 2068 controls residing in Shanghai [22] and Jiangsu province [Central Han]; Replication 2: 1785 T2D cases and 2390 controls residing in Southern China [Southern Han]) [28]. All participants self-reported as Han Chinese. T2D case was diagnosed according to the 1999 World Health Organization criteria (fasting plasma glucose level  $\geq 7.0$  mmol/l and/or 2h oral glucose tolerance test (OGTT) plasma glucose level  $\geq 11.1$  mmol/l) or with taking anti-diabetic therapies. The controls were defined as a fasting glucose level less than 6.1 mmol/l and a 2 h OGTT plasma glucose level less than 7.8 mmol/l.

### Genotyping and quality control

Genomic DNA was extracted from peripheral blood by standard phenol/chloroform-based method. In the discovery stage, genotyping was conducted by using Illumina Human 660- and 610-Quad BeadChips at the Chinese National Human Genome Center at Shanghai. Genotyping was performed according to the Infinium HD protocol from Illumina. Quality control involved exclusion of SNPs with a call rate  $< 90\%$ , a minimum allele frequency  $< 0.01$ , and/or a significant deviation from Hardy-Weinberg equilibrium (HWE) in the controls ( $P < 10^{-7}$ ). SNPs on the X, Y and mitochondrial chromosomes and copy number variation (CNV) probes were also excluded from further analysis.

In the replication stage, genotyping was conducted using 5' nuclease allelic discrimination assay (TaqMan Assay) on an ABI PRISM 7900HT Sequence Detection System following the manufacturer's protocol. In our study, the call rate ranged from 97.95% (rs3773159 in the Replication 1) to 99.33% (rs3773159 in Replication 2). There is no significant difference of SNP calling between the case and the control groups. The average consensus rate in the duplicate samples ( $n = 100$ ) was 100%.

### Statistical analysis

Identification of cryptic relatedness among individuals in the discovery stage was based on pairwise identity by state using the PLINK 1.07 software [30], after which one of the two related individuals was excluded. Population structure of the remaining sample was examined and outliers were excluded based on multidimensional scaling analysis (MDS) using PLINK 1.07 as well as principal component analysis (PCA) using EIGENSTRAT software [31]. Cochran-Armitage trend test was used to examine the association of genotype with disease phenotype and calculate odds ratio (OR) per allele. The quantile-quantile plot was employed to evaluate the overall significance of the genome-wide association results and impacts of population stratification. The genomic control inflation factor  $\lambda$  was also calculated to examine the effects of population stratification.

Replication analysis was performed by first analyzing replication samples separately and then combining them with the discovery sample set [32]. Association analysis of the combined samples was performed based on Cochran-Mantel-Haenszel tests [33]. Joint analysis was performed using PLINK under a fixed-effect model [30].

### Supporting Information

**Figure S1 Multidimensional scaling analysis (MDS) plot.** MDS plot by PLINK of the 793 cases and 806 controls shows no evident population stratification and outliers. Blue: control; pink: case. (TIF)

**Figure S2 Principal component analysis (PCA) plot.** PCA plot by EIGENSTRAT of the 793 cases and 806 controls shows no evident population stratification and outliers. Green: control; red: case. (TIF)

**Figure S3 Quantile-quantile (Q-Q) plot for the trend test. ( $\lambda = 1.08$ ).** Q-Q plot for the Cochran-Armitage trend test for 474,515 SNPs in 793 cases and 806 controls.  $\lambda = 1.08$  and minimal evidence of association due to population stratification was observed. (TIF)

**Figure S4 Manhattan plot for GWAS data.** The x-axis represents chromosomal location of 474,515 SNPs examined and the y-axis represents  $-\log_{10}$  of the P value of the Cochran-Armitage trend test under an additive model. A cutoff line was drawn at the significance threshold of  $10^{-4}$ . (TIF)

**Table S1 SNPs selected for fast-track replication.** RAF(T2D) and RAF(NC), risk allele frequency in T2D cases and controls, respectively. OR, odds ratio for risk allele. (DOC)

**Table S2 SNPs in previously reported T2D loci.** RAF(T2D) and RAF(NC), risk allele frequency in T2D cases and controls, respectively. OR, odds ratio for risk allele. \*,  $P < 0.05$ . \*\*,  $P < 0.10$ . (DOC)

### Acknowledgments

The authors thank the field workers for their contribution and the participants for their cooperation, as well as Minglan Yang, Liying Zhu, Yuanyuan Xu, Nan Liu, Xia Zhang, Qun Yan, Sheng Zheng, Xiaoyan Xie, Lijuan Li, Liyun Shen, Hongjie Qian, and Hanxiao Sun for performing the DNA preparation and experiment.

## Author Contributions

Conceived and designed the experiments: GN WH BC. Performed the experiments: XZ MX TG LX. Analyzed the data: GN XZ MX WH HW

BC. Contributed reagents/materials/analysis tools: GN DZ GC Xuejun Li YB YC YX Xiaoying Li WW WH. Wrote the paper: GN BC XZ MX.

## References

- Stumvoll M, Goldstein BJ, van Haeften TW (2005) Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 365: 1333–1346.
- Zimmet P, Alberti KG, Shaw J (2001) Global and societal implications of the diabetes epidemic. *Nature* 414: 782–787.
- Yang W, Lu J, Weng J, Jia W, Ji L, et al. (2010) Prevalence of diabetes among men and women in China. *N Engl J Med* 362: 1090–1101.
- O’Rahilly S (2009) Human genetics illuminates the paths to metabolic disease. *Nature* 462: 307–314.
- Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, et al. (2000) The common PPAR $\gamma$  Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26: 76–80.
- Gloyn AL, Weedon MN, Owen KR, Turner MJ, Knight BA, et al. (2003) Large-scale association studies of variants in genes encoding the pancreatic beta-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.
- Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38: 320–323.
- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881–885.
- Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336–1341.
- Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, 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.
- Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, et al. (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316: 1341–1345.
- Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, et al. (2007) A variant in *CDKALI* influences insulin response and risk of type 2 diabetes. *Nat Genet* 39: 770–775.
- Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 40: 638–645.
- Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, et al. (2008) Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet* 40: 1092–1097.
- Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, et al. (2008) SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet* 40: 1098–1102.
- Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, et al. (2009) Genetic variant near *IRSI* is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat Genet* 41: 1110–1115.
- Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, et al. (2009) Parental origin of sequence variants associated with complex diseases. *Nature* 462: 868–874.
- Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, et al. (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 42: 105–116.
- Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, et al. (2010) Genetic variation in gastric inhibitory polypeptide receptor (GIPR) impacts the glucose and insulin responses to an oral glucose challenge. *Nat Genet* 42: 142–148.
- Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, et al. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 42: 579–589.
- Xiang J, Li XY, Xu M, Hong J, Huang Y, et al. (2008) Zinc transporter-8 gene (*SLC30A8*) is associated with type 2 diabetes in Chinese. *J Clin Endocrinol Metab* 93: 4107–4112.
- Xu M, Bi Y, Xu Y, Yu B, Huang Y, et al. (2010) Combined effects of 19 common variations on type 2 diabetes in Chinese: results from two community-based studies. *PLoS One* 5: e14022.
- Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, et al. (2010) Genome-wide association studies in diverse populations. *Nat Rev Genet* 11: 356–366.
- McCarthy MI (2008) Casting a wider net for diabetes susceptibility genes. *Nat Genet* 40: 1039–1040.
- Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, et al. (2010) A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. *PLoS Genet* 6: e1000847.
- Shu XO, Long J, Cai Q, Qi L, Xiang YB, et al. (2010) Identification of new genetic risk variants for type 2 diabetes. *PLoS Genet* 6: e1001127.
- Yamauchi T, Hara K, Maeda S, Yasuda K, Takahashi A, et al. (2010) A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at *UBE2E2* and *C2CD4A-C2CD4B*. *Nat Genet* 42: 864–868.
- Xu S, Yin X, Li S, Jin W, Lou H, et al. (2009) Genomic dissection of population substructure of Han Chinese and its implication in association studies. *Am J Hum Genet* 85: 762–774.
- Xu M, Li XY, Wang JG, Wang XJ, Huang Y, Cheng Q, et al. (2009) Retinol-binding protein 4 is associated with impaired glucose regulation and microalbuminuria in a Chinese population. *Diabetologia* 52: 1511–9.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
- Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 38: 209–213.
- Mantel N, Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22: 719–748.