

# SCIENTIFIC REPORTS



OPEN

## *TOX* and *CDKN2A/B* Gene Polymorphisms Are Associated with Type 2 Diabetes in Han Chinese

Received: 27 November 2014

Accepted: 08 June 2015

Published: 03 July 2015

Fengjiang Wei<sup>1,\*,</sup> Chunyou Cai<sup>1,\*,</sup> Shuzhi Feng<sup>2,\*,</sup> Jia Lv<sup>1,</sup> Shen Li<sup>1,</sup> Baocheng Chang<sup>3,</sup> Hong Zhang<sup>4,</sup> Wentao Shi<sup>1,</sup> Hongling Han<sup>2,</sup> Chao Ling<sup>1,</sup> Ping Yu<sup>1,</sup> Yongjun Chen<sup>4,</sup> Ning Sun<sup>2,</sup> Jianli Tian<sup>2,</sup> Hongxiao Jiao<sup>1,</sup> Fuhua Yang<sup>1,</sup> Mingshan Li<sup>1,</sup> Yuhua Wang<sup>1,</sup> Lei Zou<sup>1,</sup> Long Su<sup>4,</sup> Jingbo Li<sup>5,</sup> Ran Li<sup>2,</sup> Huina Qiu<sup>5,</sup> Jingmin Shi<sup>1,</sup> Shiyong Liu<sup>2,</sup> Mingqin Chang<sup>1,</sup> Jingna Lin<sup>5,</sup> Liming Chen<sup>3</sup> & Wei-Dong Li<sup>1</sup>

To study associations between type 2 diabetes (T2DM) candidate genes and microvascular complications of diabetes (MVCDs), we performed case-control association studies for both T2DM and MVCDs in Han Chinese subjects. We recruited 1,939 unrelated Han Chinese T2DM patients and 918 individuals with normal blood glucose levels as nondiabetic controls. Among T2DM patients, 1116 have MVCDs, 266 have a history of T2DM of >10 years but never developed MVCDs. Eighty-two single-nucleotide polymorphisms (SNPs) in 54 candidate genes were genotyped. Discrete association studies were performed by the PLINK program for T2DM and MVCDs. Significant associations were found among candidate gene SNPs and T2DM, including rs1526167 of the *TOX* gene (allele A,  $P=2.85 \times 10^{-9}$ , OR=1.44). The SNP rs10811661 of the *CDKN2A/B* gene was also associated with T2DM (allele T,  $P=4.09 \times 10^{-7}$ , OR=1.36). When we used control patients with >10 years of T2DM history without MVCD, we found that the G allele of SNP rs1526167 of the *TOX* gene was associated with MVCD (nominal  $P=4.33 \times 10^{-4}$ ). In our study, significant associations were found between *TOX* and *CDKN2A/B* gene SNPs and T2DM. The *TOX* polymorphism might account for the higher risk of T2DM and the lower risk of MVCDs in the Han Chinese population.

The prevalence of type 2 diabetes (T2DM) has increased dramatically in China in recent years<sup>1,2</sup>. Many T2DM genes found in European populations have been replicated in Han Chinese; however, many top candidate genes (including *TCF7L2*) have shown only moderate associations with T2DM in East Asian populations. On the other hand, many Asian-specific T2DM genes have been identified in recent studies<sup>3</sup>.

With the remarkable rise worldwide in the prevalence of diabetes, an increase in patients suffering from microvascular complications of diabetes (MVCDs) will be inevitable. Diabetic nephropathy (DN) and diabetic retinopathy (DR), two common MVCDs, are leading causes of end-stage renal failure and blindness in diabetes patients<sup>4,5</sup>. Epidemiology data for MVCDs are relatively limited in China, although a recent study showed a much lower prevalence of DR in the Beijing area than in European populations<sup>6</sup>.

<sup>1</sup>Research Center of Basic Medical Sciences, Tianjin Medical University, Tianjin, 300070, China. <sup>2</sup>Tianjin General Hospital, Tianjin Medical University, Tianjin, 300052, China. <sup>3</sup>Metabolic Diseases Hospital, Tianjin Medical University, Tianjin, 300070, China. <sup>4</sup>Eye Hospital, Tianjin Medical University, Tianjin, 300384, China. <sup>5</sup>Tianjin People's Hospital, Department of Endocrinology, Tianjin, 300191, China. <sup>\*</sup>These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.L. (email: linjingna6@yahoo.com) or L.C. (email: xfx22081@vip.163.com) or W.-D.L. (email: liweidong98@tjmu.edu.cn)

SNP	Chr	Position	Gene	Risk allele	Risk allele frequencies	P	OR (95% CI)
rs10811661	9	22134094	<i>CDKN2A/B</i>	T	0.56	$4.09 \times 10^{-7*}$	1.364 (1.209~1.538)
rs1526167	8	59702355	<i>TOX</i>	A	0.46	$2.85 \times 10^{-9*}$	1.440 (1.276~1.624)
rs4402960	3	185511687	<i>IGF2BP2</i>	T	0.25	$7.56 \times 10^{-4}$	1.352 (1.134~1.612)
rs6856526	4	61057462	<i>LPHN3</i>	C	0.93	$1.33 \times 10^{-3}$	1.522 (1.176~1.970)
rs13266634	8	118184783	<i>SLC30A8</i>	C	0.64	$4.56 \times 10^{-3}$	1.282 (1.080~1.532)
rs7756992	6	20679709	<i>CDKAL1</i>	G	0.54	$8.21 \times 10^{-4}$	1.246 (1.095~1.418)

**Table 1. Association analyses for T2DM (1939 cases, 918 non-T2DM controls, nominal  $P < 5 \times 10^{-3}$ ).**

\*Significant for genome-wide association ( $P < 5 \times 10^{-8}$ ). \*Significant after multiple tests correction.

Although previous studies have shown that the duration of diabetes and plasma glucose levels are MVCD risk factors<sup>7,8</sup>, some diabetes patients with severe hyperglycemia never develop MVCDs. In contrast, many diabetes patients with well-controlled blood glucose suffer from MVCDs<sup>9</sup>. Many studies have found family aggregation of MVCDs and significantly increased risk in siblings, suggested that genetic factors play an important role in the etiology of MVCDs<sup>10–14</sup>. Although many association studies have identified genes related to MVCDs<sup>15</sup>, the search for MVCD susceptibility genes was less successful than for T2DM susceptibility genes, partially because of un-diagnosed MVCDs and lack of appropriate normal controls.

It is possible that Han Chinese-specific gene variants are associated with T2DM and/or MVCDs. In this study, we aimed to find T2DM and MVCD candidate genes in Han Chinese T2DM patients by studying associations between single-nucleotide polymorphisms (SNPs) of candidate genes and T2DM/MVCDs.

## Results

Using 1,939 T2DM patients as cases and 918 nondiabetic individuals (>57 years old, with normal blood glucose) as controls, analyses for T2DM showed association between candidate gene SNPs and T2DM (Table 1), including rs10811661 of the *CDKN2A/B* gene (allele T,  $P = 4.09 \times 10^{-7}$ , OR = 1.364, 95% CI = 1.209–1.538), rs1526167 of the *TOX* gene (allele A,  $P = 2.85 \times 10^{-9}$ , OR = 1.440, 95% CI = 1.276–1.624), rs4402960 of the *IGF2BP2* gene (allele T,  $P = 7.56 \times 10^{-4}$ , OR = 1.352, 95% CI = 1.134–1.612), rs6856526 of the *LPHN3* gene (allele C,  $P = 1.33 \times 10^{-3}$ , OR = 1.522, 95% CI = 1.176–1.970), rs13266634 of the *SLC30A8* gene (allele C,  $P = 4.56 \times 10^{-3}$ , OR = 1.282, 95% CI = 1.080–1.532), and rs7756992 of the *CDKAL1* gene (allele G,  $P = 8.21 \times 10^{-4}$ , OR = 1.246, 95% CI = 1.095–1.418). The *TOX* gene SNP rs1526167 association reached genome-wide association level ( $P < 5 \times 10^{-8}$ ), while the *CDKN2A/B* SNP rs10811661 was significant after Bonferroni correlation ( $P < 1.22 \times 10^{-4}$  for 82 SNPs and 5 binary traits,  $6.1 \times 10^{-4}$  for 82 SNPs and 1 binary trait).

In MVCD association studies, DN, DR, proliferative retinopathy (PDR), and MVCD patients were selected as cases, and patients with >10 years T2DM history, without DR or DN, were chosen as controls. Association analyses of dichotomous variables showed that rs1526167 of the *TOX* gene was associated with DN ( $P = 0.0011$ , OR = 1.470, 95% CI = 1.166–1.854), DR ( $P = 0.0082$ , OR = 1.412, 95% CI = 1.093–1.825), and MVCD ( $P = 4.33 \times 10^{-4}$ , OR = 1.498, 95% CI = 1.195–1.878). However, the at-risk allele for MVCDs was G, rather than A for T2DM. The T allele of rs10811661 of the *CDKN2A/B* gene was associated with DR ( $P = 0.037$ , OR = 1.314, 95% CI = 1.016–1.698), PDR ( $P = 0.026$ , OR = 1.708, 95% CI = 1.063–2.743), and MVCD ( $P = 0.025$ , OR = 1.292, 95% CI = 1.032–1.617). The G allele of rs4402960 of the *IGF2BP2* gene was associated with DN ( $P = 0.0092$ , OR = 1.499, 95% CI = 1.104–2.034) and MVCD ( $P = 0.0094$ , OR = 1.481, 95% CI = 1.100–1.995). The T allele of rs12102171 of the *SMAD3* gene was associated with DR and MVCD ( $P = 0.027$ , and 0.045, respectively). The A allele of the *ESR1* gene was associated with DR and PDR ( $P = 0.041$  and 0.033, respectively) (Table 2). None of these nominal associations for MVCDs remained significant after Bonferroni corrections.

Four SNPs of the *TOX* gene were genotyped (Table 3), rs1526167 was not in linkage disequilibrium with other 3 SNPs (Supplement Table 2, Supplement Figures 1 and 2). The SNP rs17304270 of the *TOX* gene gave marginal association with diabetic nephropathy: allele “A”, OR = 1.514 (1.003–2.285),  $P = 0.047$ .

## Discussion

To date, more than 30 genes have been identified as reaching the genome-wide significance threshold ( $P < 5 \times 10^{-8}$ ) for T2DM; 10 of these were replicated in the Han Chinese population<sup>16</sup>. More than 60% of T2DM genes found in East Asian genome-wide association studies (GWAS) were replicated in Han Chinese<sup>17</sup>, compared with approximately 30% of European population T2DM genes. Compared with Caucasians, the Han Chinese and East Asian populations are more insulin resistant, are more centrally obese, and develop T2DM more readily even with moderate increases in body mass index<sup>18–20</sup>. Thus,

SNP	rs4402960	rs1526167	rs12102171	rs722208	rs10811661
chromosome	3	8	15	6	9
position	185511687	59702355	67425033	152322885	22134094
Gene	<i>IGF2BP2</i>	<i>TOX</i>	<i>SMAD3</i>	<i>ESR1</i>	<i>CDKN2A/B</i>
Risk allele	G	G	T	A	T
DN (N = 836)					
Risk allele frequencies	0.72	0.48			
P	0.0092	0.0011			
OR (95% CI)	1.499 (1.104 ~ 2.034)	1.470 (1.166 ~ 1.854)			
DR (N = 504)					
Risk allele frequencies		0.46	0.31	0.46	0.58
P		0.0082	0.027	0.041	0.037
OR (95% CI)		1.412 (1.093 ~ 1.825)	1.471 (1.043 ~ 2.072)	1.305 (1.011 ~ 1.686)	1.314 (1.016 ~ 1.698)
PDR (N = 106)					
Risk allele frequencies				0.44	0.57
P				0.033	0.026
OR (95% CI)				1.632 (1.038 ~ 2.566)	1.708 (1.063 ~ 2.743)
MVCD (N = 1116)					
Risk allele frequencies	0.72	0.49	0.31		0.59
P	0.0094	$4.33 \times 10^{-4}$	0.045		0.025
OR (95% CI)	1.481 (1.100 ~ 1.995)	1.498 (1.195 ~ 1.878)	1.376 (1.007 ~ 1.879)		1.292 (1.032 ~ 1.617)

**Table 2. Candidate gene association studies for MVCD (266 non-MVCD T2DM patients as controls).** Only results with nominal  $P < 0.05$  are shown.

seeking T2DM genes in Han Chinese would help us to better understanding the global genetic background of T2DM.

In this study, as candidate genes we chose 1) genes related to T2DM, obesity, or insulin resistance found by previous GWAS and 2) genes related to glucose and lipid metabolism, insulin secretion, or MVCD, such as the HIF1 $\alpha$ /-VEGF pathway.

A significant association ( $P = 2.85 \times 10^{-9}$ ) between a *TOX* gene SNP and T2DM was identified for the first time by this study. In addition, four previous reported T2DM genes, *CDKN2A/B*, *IGF2BP2*, *SLC30A8*, and *CDKAL1*, showed moderate to strong associations with T2DM. We did not replicate associations with several well-known T2DM genes, including *TCF7L2*, *FTO*, *IRS1*, and *KCNQ1*, in our Han Chinese population (In our data set, the *FTO* gene SNPs yielded some associations with insulin resistance related phenotypes in quantitative analyses. Data not shown). Potential reasons for this failed replication may include genetic heterogeneity, low allele frequencies of tested SNPs in Han Chinese (e.g., the MAF of *TCF7L2* gene SNP rs7903146 was 0.046), less covered genes (i.e., too few SNPs were genotyped for certain candidate genes), and the relatively small sample size. We performed a gene-specific power calculations for genes *IRS1*, *TCF7L2*, *KCNQ1*, and *FTO* using real minor allele frequencies (MAF) in our data set (Supplement Table 1). Except *TCF7L2*, we have adequate power ( $> 80\%$ ) to detect associations at the type I error rate ( $\alpha$ ) = 0.05 when genotype relative risk (GRR)  $> 1.4$ . We have limited power when  $GRR \leq 1.2$ , it could be the reason of some failed replications for T2DM candidate genes in our study.

Compared with studies of T2DM, GWAS and association studies of MVCD are relatively limited<sup>21-26</sup>. So far about half a dozen of GWAS were performed for MVCD (including 3 in East Asian populations), although none of the MVCD associations reached the genome-wide significance threshold of  $P < 5 \times 10^{-8}$ . In order to subtract the influence of T2DM, we selected as MVCD controls subjects with  $> 10$  years history of T2DM who never developed DN or DR. We had only 106 PDR patients in the study, but we still achieved moderate nominal  $P$  values for *CDKN2A/B* and *ESR1* gene SNPs. Recently, Sheu *et al.*<sup>22</sup> performed a GWAS for DR in a Chinese population, several loci was associated with DR, although no genome wide association with  $P < 1 \times 10^{-7}$  was reached. We have not gotten a chance to test their findings in our population, on the other hand, the *TOX* and *CDKN2A/B* polymorphisms were not among their top associations. We also tested candidate genes for DN that identified by GWAS in African Americans<sup>21,27</sup>, including *MYH9*, *SFI1*, and *LIMK2*, but no association was found for MVCD. Given the sample size of our non-MVCD T2DM controls (266), we only have moderate power for DN and DR association studies. For PDR, the detection power was very limited. On the other hand, we have *MYH9* and *SFI1* gene well covered in our study (8 SNPs for *MYH9* and 3 for *SFI1*,  $D' > 0.6$ ), we have

SNP	Gene	Chr	HWE ( <i>P</i> )	MAF*		
				CHB	CEU	Global
rs7546903	<i>CAMTA1</i>	1	0.603	0.463	0.226	0.368
rs1801133	<i>MTHFR</i>	1	0.236	0.439	0.310	0.325
rs6427665	<i>NOS1AP</i>	1	0.942	0.455	0.233	0.378
rs2661812	<i>NOS1AP</i>	1	0.596	0.475	0.500	0.445
rs16867321	<i>UBE2E3</i>	2	1.000	0.415	0.200	0.271
rs62183937	<i>ABI2</i>	2	1.000	0.475	0.125	0.258
rs11675251	<i>ABI2</i>	2	0.303	0.171	0.482	0.381
rs3731652	<i>ABI2</i>	2	0.666	0.433	0.158	0.375
rs1376877	<i>ABI2</i>	2	0.177	0.171	0.455	0.383
rs11677793	<i>SPAG16</i>	2	0.992	0.200	0.456	0.284
rs7578326	<i>IRS1</i>	2	0.512	0.125	0.350	0.304
rs1678607	<i>VHL</i>	3	0.768	0.111	0.125	0.208
rs13081389	<i>PPAR<math>\gamma</math></i>	3	0.438	0.022	0.042	0.034
rs35747495	<i>PCAF</i>	3	0.979	0.325	0.300	0.260
rs2929402	<i>PCAF</i>	3	0.227	0.463	0.372	0.419
rs1986917	<i>PCAF</i>	3	0.987	0.433	0.442	0.389
rs4402960	<i>IGF2BP2</i>	3	0.188	0.256	0.280	0.343
rs13129697	<i>SLC2A9</i>	4	0.920	0.439	0.292	0.423
rs1014290	<i>SLC2A9</i>	4	0.892	0.363	0.257	0.308
rs6856526	<i>LPHN3</i>	4	0.484	0.073	0.009	0.129
rs2231142	<i>ABCG2</i>	4	0.989	0.293	0.111	0.139
rs10946398	<i>CDKAL1</i>	6	0.851	0.439	0.336	0.408
rs7756992	<i>CDKAL1</i>	6	0.312	0.488	0.279	0.405
rs1165196	<i>SLC17A1</i>	6	0.205	0.232	0.451	0.260
rs881858	<i>VEGFA</i>	6	0.045	0.189	0.292	0.346
rs9395706	<i>PKHD1</i>	6	0.989	0.476	0.128	0.296
rs722208	<i>ESR1</i>	6	0.583	0.500	0.246	0.412
rs4880	<i>SOD2</i>	6	0.713	0.111	0.300	0.188
rs1581498	<i>SNORD93</i>	7	0.362	0.400	0.467	0.378
rs1799884	<i>GCK</i>	7	0.084	0.171	0.195	0.188
rs705382	<i>PON1</i>	7	0.755	0.415	0.336	0.472
rs1007311	<i>NOS3</i>	7	0.587	0.308	0.500	0.437
rs768403	<i>GBX1</i>	7	0.877	0.463	0.398	0.476
rs7805834	<i>NUB1</i>	7	0.694	0.073	0.102	0.140
rs446886	<i>NUB1</i>	7	0.151	0.073	0.310	0.298
rs386956	<i>NUB1</i>	7	0.972	0.488	0.319	0.478
rs1526167	<i>TOX</i>	8	0.445	0.478	0.467	0.478
rs2726557	<i>TOX</i>	8	0.254	0.427	0.327	0.459
rs11777927	<i>TOX</i>	8	0.757	0.356	0.267	0.356
rs17304270	<i>TOX</i>	8	0.130	0.061	0.288	0.390
rs13266634	<i>SLC30A8</i>	8	0.930	0.476	0.239	0.282
rs10811661	<i>CDKN2A/B</i>	9	0.764	0.415	0.199	0.206
rs3758391	<i>SIRT1</i>	10	0.965	0.195	0.270	0.473
rs7923837	<i>HHEX</i>	10	1.000	0.244	0.367	0.427
rs7903146	<i>TCF7L2</i>	10	0.236	0.024	0.279	0.218
rs2237892	<i>KCNQ1</i>	11	0.287	0.317	0.075	0.170
rs2166706	<i>MTNR1B</i>	11	0.658	0.317	0.389	0.472
rs189037	<i>ATM</i>	11	0.997	0.389	0.485	0.485

Continued

SNP	Gene	Chr	HWE ( <i>P</i> )	MAF*		
				CHB	CEU	Global
rs7312112	<i>IGF1</i>	12	0.304	0.500	0.385	0.462
rs2241220	<i>ACACB</i>	12	0.996	0.341	0.142	0.109
rs11067076	<i>TBX5</i>	12	0.586	0.037	0.257	0.193
rs11067083	<i>TBX5</i>	12	0.545	0.064	0.242	0.199
rs371276	<i>SLITRK5</i>	13	0.615	0.463	0.025	0.244
rs409762	<i>SLITRK5</i>	13	0.832	0.488	0.013	0.244
rs2301113	<i>HIF1A</i>	14	0.925	0.317	0.173	0.428
rs11624704	<i>NRXN3</i>	14	0.492	0.049	0.137	0.115
rs1498506	<i>SMAD3</i>	15	0.840	0.433	0.475	0.455
rs12102171	<i>SMAD3</i>	15	0.353	0.341	0.195	0.276
rs17818920	<i>FTO</i>	16	1.000	0.183	0.250	0.258
rs1876942	<i>FTO</i>	16	0.309	0.341	0.425	0.479
rs708254	<i>FTO</i>	16	0.438	0.350	0.420	0.390
rs2239359	<i>FANCA</i>	16	0.970	0.207	0.416	0.393
rs7190823	<i>FANCA</i>	16	0.488	0.024	0.416	0.343
rs4353	<i>ACE</i>	17	0.953	0.317	0.495	0.498
rs17782313	<i>MC4R</i>	18	0.489	0.232	0.265	0.221
rs8109627	<i>CCDC97</i>	19	0.928	0.390	0.235	0.341
rs4814615	<i>PCSK2</i>	20	1.000	0.488	0.128	0.293
rs3746876	<i>KCNJ15</i>	21	0.906	0.110	0.004	0.046
rs2106294	<i>LIMK2</i>	22	0.824	0.073	0.305	0.178
rs5749286	<i>SFI1</i>	22	0.875	0.378	0.283	0.271
rs5753669	<i>SFI1</i>	22	0.497	0.378	0.283	0.271
rs2295251	<i>SFI1</i>	22	0.652	0.451	0.270	0.404
rs735853	<i>MYH9</i>	22	0.991	0.110	0.477	0.273
rs875726	<i>MYH9</i>	22	0.456	0.305	0.296	0.449
rs2009930	<i>MYH9</i>	22	0.377	0.275	0.283	0.450
rs2239782	<i>MYH9</i>	22	0.067	0.317	0.239	0.349
rs3752462	<i>MYH9</i>	22	0.320	0.305	0.332	0.453
rs2269532	<i>MYH9</i>	22	0.881	0.267	0.358	0.390
rs2071731	<i>MYH9</i>	22	0.999	0.280	0.367	0.423
rs739097	<i>MYH9</i>	22	0.876	0.268	0.456	0.495
rs2285094	<i>PDGFB</i>	22	0.549	0.183	1.000	0.328
rs738409	<i>PNPLA3</i>	22	0.804	0.344	0.233	0.284

**Table 3. Candidate genes and SNPs genotyped in this study.** \*MAF: minor allele frequencies, taken from dbSNP. CHB, Han Chinese; CEU, European American.

>80% power to detect association for MVCD, DN, and DR when GRRs were more than 1.40, 1.42, 1.47, respectively ( $\alpha = 0.05$ , dominant model). Compare with T2DM, our power for detection of MVCD associations was moderate, mainly because of the limited sample size of the control group. We cannot rule out the *MYH9* and *SFI1* associations that found in African Americans, although population heterogeneity could be a major reason of the difference between Han Chinese and African Americans.

In this study, SNPs in *TOX* and *CDKN2A/B* genes yielded the most significant associations for both T2DM and MVCD. Interestingly, the A allele of the *TOX* gene SNP rs1526167 was the “at-risk” allele in T2DM and the “protection” allele for MVCD. Recent research carried out in Beijing (very close to Tianjin, where our subjects were collected) showed that the prevalence of NPDR was 18.6% in Han Chinese T2DM patients<sup>6</sup>, compared with 28.5% in U.S. T2DM patients, based on NHANES 2005–2008 data<sup>28</sup>. Prevalence of DR was quite different among U.S. ethnic groups and was higher in African Americans than in European Americans<sup>28</sup>. To determine whether the “protection” allele of the *TOX* gene accounts for the lower DR prevalence in Han Chinese MVCD patients, studies with larger sample sizes are needed.

The *CDKN2A/B* gene is located in the chromosome 9p21 region, which has been highlighted as the strongest genetic susceptibility locus for cardiovascular disease (CVD)<sup>29,30</sup> and linked to other conditions such as T2DM<sup>30,31</sup>, Alzheimer's disease<sup>32</sup>, glaucoma<sup>33</sup>, and periodontitis<sup>34</sup>. Interestingly, the region of the chromosome associated with CVD and diabetes was previously considered a gene desert. Previous studies found that rs10811661 has a strong correlation with T2DM in French<sup>35</sup>, Japanese<sup>36</sup>, and Chinese<sup>37</sup> populations. The SNP rs10811661 locates ~100 kb upstream of *CDKN2A/B*, which has been shown to associate with downregulation of antisense noncoding RNA in *INK4* locus (*ANRIL*) expression<sup>38</sup>. The *ANRIL* methylates histone H3K27 by interacting with polycomb proteins, therefore suppresses the expression of *INK4a* (i.e. *CDKN2A*)<sup>39</sup>. Recently, a large sample sized candidate gene association study for T2DM was carried out in Chinese population<sup>40</sup>, 8 T2DM related genes from previous GWASs were replicated. The *CDKN2A/B* region SNP, rs10811661, yielded the most significant association ( $P = 1.11 \times 10^{-8}$ ). The SNP was also associated with many quantitative glycemic traits.

At present, more studies have been done on the correlation between *CDKN2A/B* and the pathogenesis of T2DM than on the relationship between *CDKN2A/B* and MVCD<sup>41–43</sup>. In the present study, we found associations with both T2DM and MVCD for rs10811661, and the allele T contributed to the increased risk for both diseases. Although associations on the *CDKN2A/B* (*ANRIL*) locus were well documented, the mechanism by which this locus affects susceptibility for T2DM and MVCD remains to be investigated.

The *TOX* gene association for T2DM was first discovered in the present study. The *TOX* gene, a protein-coding gene located in human chromosome 8, is one of the *TOX* high-mobility-group proteins. In the present study, we tested 4 SNPs of the *TOX* gene, rs1526167, rs2726557, rs1177927, and rs17304270. The SNP rs1526167 was not in linkage disequilibrium with other 3 SNPs (Supplement Table 2), it could be the reason for the lack of association of T2DM with those 3 SNPs. The SNP rs17304270, however, was nominally associated with diabetic nephropathy.

The history of the SNP rs1526167 was a little complicated. The SNP rs1526167 located 15Kb downstream of the *TOX* gene and 1Kb upstream of an unknown function mRNA *DL491802*. The SNP was first identified by Perlegen Sciences. At the time of the International HapMap Project, that SNP was not included. It could be the reason why rs1526167 was not included in most main stream genome-wide SNP genotyping panels. We checked the LD pattern (in  $r^2$ ) of the chr 8: 59,830,000–59,920,000 region in the HapMap for both Chinese (CHB) and Caucasian (CEU) populations (Supplement Figures 1 and 2): the SNP rs1526167 was located between rs10090702 and rs2726588 (indicated by arrows), and it was located in a separate haplotype block, not in LD with SNPs in the *TOX* gene coding region and introns.

Although there was no rs1526167 association reported for T2DM, there were some associations found on that SNP for obesity and metabolic syndrome related phenotypes. In a published US patent application (Pub No: US 2006/0177847 A1, Pub Date: Aug 10, 2006), Cox *et al.* found the *TOX* gene polymorphism and other 27 DNA sequence variations were associated with Olanzapine treatment emergent weight gain and “metabolic syndrome” in a 1.7 million SNPs genome association study. After checking the DNA sequence provided by the patent application, rs1526167 was among the candidate gene SNPs to screen Olanzapine treatment emergent weight gain and other related traits.

The SNP rs1526167 was associated with obesity in European American extremely obese trios by a transmission disequilibrium test (TDT):  $P = 2.2 \times 10^{-5}$ ; all probands had BMI >35 kg/m<sup>2</sup>, 428 European American trios (Price *et al.*, unpublished data).

Our results showed that rs1526167 was associated with both T2DM and MVCD, although the risk allele was different for those two conditions. The biological connections between *TOX* and T2DM are poorly understood, although *TOX* gene polymorphisms are associated with insulin resistance traits in both Han Chinese (present study) and European Americans.

Since we have not genotyped all reported “positive” SNPs for our candidate genes, we failed to replicate many well established associations. However, the main purpose of this study is not to exclude certain genes from the T2DM/MVCD candidates. We have selected most of our SNPs based on minor allele frequencies (MAF) in Han Chinese, although several SNPs that with higher MAFs in Caucasians were genotyped to verify our previous findings.

In our study, we employed a well-phenotyped, ethnically homogenous population of diabetic subjects, although the number of non-MVCD T2DM controls was relatively limited. A U.S. NHANES epidemiological study showed that prevalence of MVCDs reaches its peak 15 years after onset of T2DM, with almost no new MVCDs developing in individuals with >15 years of T2DM history<sup>28</sup>. In the present study, we considered individuals with a history of T2DM >10 years, without MVCDs, to be unlikely to carry MVCD susceptibility genes. In our study, the duration of T2DM history in non-MVCD controls was  $17.0 \pm 5.2$  years. We therefore selected these individuals as non-MVCD controls. More controls need to be recruited for association studies with larger power, and much more genotyping needs to be done in the subjects to better understand these relationships between SNPs and MVCDs.

## Materials and Methods

**Study population and design.** We recruited 1,939 T2DM diabetic patients and 918 individuals >57 years old with normal blood glucose levels as nondiabetic controls. We collected older subjects for controls since the late onset of T2DM. All the subjects were unrelated Han Chinese collected from the Metabolic Disease Hospital of Tianjin Medical University, General Hospital of Tianjin Medical University, Tianjin



	T2DM	Non-diabetic controls	DN	DR	PDR	MVCD	Non-MVCD T2DM controls
N	1,939	918	836	504	106	1,116	266
Sex (male/female)	1,040/899	740/178	482/354	257/247	44/62	605/511	124/142
Age (yr, mean $\pm$ SD)	58.5 $\pm$ 12.0	72.4 $\pm$ 9.1	58.8 $\pm$ 11.1	57.8 $\pm$ 11.1	54.6 $\pm$ 11.1	58.7 $\pm$ 11.1	63.4 $\pm$ 9.3
T2DM duration (yr, mean $\pm$ SD)	10.3 $\pm$ 7.7	N/A	11.3 $\pm$ 7.6	12.6 $\pm$ 7.6	13.1 $\pm$ 8.2	11.6 $\pm$ 7.7	17.0 $\pm$ 5.2

**Table 4. Basic characteristics of T2DM and MVCD cases and controls.**

People's Hospital, and Eye Hospital of Tianjin Medical University. We examined the patients for DR and DN: 836 patients had DN, 398 had nonproliferative DR (NPDR), 106 had proliferative DR (PDR), and 504 had DR; 224 patients had both DN and DR (612 and 280 subjects only had DN or DR, respectively). Overall, 1,116 patients had MVCD (DN or DR). As non-MVCD controls, we used 266 patients with a history of T2DM for >10 years who never developed DR or DN; the average duration of T2DM history in non-MVCD controls was 17.0  $\pm$  5.2 years (Table 4).

All subjects gave written informed consent prior to this study, and the protocol was approved by the Committee on Studies Involving Human Beings at Tianjin Medical University. The study was carried out in accordance with the approved guidelines.

We collected the patient's general information and clinical characteristics, including gender, age, height, weight, biochemistry and lipid profiles, and fasting plasma glucose. Patients were examined for DR and DN. All phenotypes were documented in a Filemaker Pro database. Table 4 presents the basic characteristics of the study population.

**Diabetic retinopathy assessment.** All patients received a professional fundus examination and fundus photography; the results were checked and graded by two ophthalmologists at the Eye Hospital of Tianjin Medical University. Level of retinopathy was defined according to a new international classification of DR<sup>44</sup> developed by the 29th International Congress of Ophthalmology in 2002. This classification comprises five levels: non-DR, mild nonproliferative DR, moderate nonproliferative retinopathy, severe nonproliferative retinopathy and proliferative retinopathy (PDR). If the levels of DR were inconsistent for the two eyes, the worse eye was recorded for the patient. We used as cases both all DR patients and the subset of PDR patients for separate analyses (Table 4).

**Diabetic nephropathy measurement.** Microalbumin excretion rates were measured for each patient. The diagnostic criteria for DN includes a history of diabetes and microalbumin/creatinine >300 mg/g, or diagnosis by the renal biopsy. All patients with primary glomerular disease and other secondary glomerular diseases were excluded.

**Candidate gene selection and genotyping.** As candidate genes for case-control association studies, we selected 1) genes associated with T2DM, obesity, MVCD, or insulin resistance in our<sup>45</sup> and other previous candidate gene associations and genome wide association studies (GWAS)<sup>21,46,47</sup>; and 2) genes in biological pathways related to the development of MVCD, glucose and lipid metabolism, or insulin secretion in diabetes. We selected 82 SNPs in 54 candidate genes (Table 3). Minor allele frequencies (MAFs) of Han Chinese, European American, and global populations were taken from dbSNP (<http://www.ncbi.nlm.nih.gov/snp/>). For previously reported associations, we selected SNPs with the most significant association rather than genotyping the whole gene. For less studied genes, multiple SNPs were chosen based on the linkage disequilibrium pattern of the gene ( $D' > 0.6$ ). In this study, we have not genotyped all tagged SNPs in certain candidate genes due to limited resources. We also performed gene-specific power calculations for *IRS1*, *TCF7L2*, *KCNQ1*, and *FTO* genes based on real minor allele frequencies (MAF) in our data set (Supplement Table 1). Linkage disequilibrium among candidate gene SNPs (in  $D'$  and  $r^2$ ) was calculated by Haploview<sup>48</sup>, results were shown as Supplement Table 2. Linkage disequilibrium (LD) patterns (in  $r^2$ ) of the *TOX* gene region SNPs were shown as Supplement Figure 1 (Chinese, CHB) and Supplement Figure 2 (Caucasian, CEU).

Genomic DNA samples were extracted from peripheral whole blood samples using the high-salt method. Genotyping was performed by primer extension of multiplex products with detection by matrix-assisted laser desorption time-of-flight mass spectrometry.

**Association Studies.** The Hardy-Weinberg equilibrium (HWE) test was performed before the association analysis (Table 3). Statistical analyses for phenotypes were performed by SPSS, version 17.0. The allelic frequencies between the case group and the control group were compared by chi tests using PLINK<sup>49</sup>, and odds ratios (ORs) with 95% confidence intervals (CIs) are presented.

Association studies were carried out in two stages: 1) using 1,939 T2DM patients as cases and 918 nondiabetic individuals as controls, to test associations for T2DM; and 2) using 1,116 MVCD patients as cases and 266 patients with a history of T2DM for >10 years, but without MVCDs, as controls, to test associations for MVCDs. Association studies were performed separately for DN, DR, PDR, and MVCD.

## Conclusions

In summary, our case-control studies suggest that *TOX* and *CDKN2A/B* gene SNPs are associated with T2DM, DN, DR, and MVCD in Han Chinese. A large prospective study is needed to confirm these associations in Han Chinese. A better understanding of genetic factors predisposing individuals to diabetic complications would help identify diabetic patients at risk and also to reveal the pathogenesis of MVCD.

## References

- Xu, Y. *et al.* Prevalence and control of diabetes in Chinese adults. *JAMA* **310**, 948–959, doi: 10.1001/jama.2013.168118 (2013).
- Yang, W. *et al.* Prevalence of diabetes among men and women in China. *N Engl J Med* **362**, 1090–1101, doi: 10.1056/NEJMoa0908292 (2010).
- Cho, Y. S. *et al.* Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat Genet* **44**, 67–72, doi: 10.1038/ng.1019 (2012).
- Martini, S., Eichinger, F., Nair, V. & Kretzler, M. Defining human diabetic nephropathy on the molecular level: integration of transcriptomic profiles with biological knowledge. *Rev Endocr Metab Disord* **9**, 267–274, doi: 10.1007/s11154-008-9103-3 (2008).
- Thylefors, B., Negrel, A. D., Pararajasegaram, R. & Dadzie, K. Y. Global data on blindness. *Bull World Health Organ* **73**, 115–121 (1995).
- Yuan, M. X. *et al.* Low prevalence of diabetic retinopathy in a Chinese population. *Diabetes care* **35**, e61, doi: 10.2337/dc12-0600 (2012).
- Freeman, M. S., Mansfield, M. W., Barrett, J. H. & Grant, P. J. Heritability of features of the insulin resistance syndrome in a community-based study of healthy families. *Diabet Med* **19**, 994–999 (2002).
- Stratton, I. M. *et al.* Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ* **321**, 405–412 (2000).
- Hietala, K., Forsblom, C., Summanen, P. & Groop, P. H. Heritability of proliferative diabetic retinopathy. *Diabetes* **57**, 2176–2180, doi: 10.2337/db07-1495 (2008).
- Quinn, M., Angelico, M. C., Warram, J. H. & Krolewski, A. S. Familial factors determine the development of diabetic nephropathy in patients with IDDM. *Diabetologia* **39**, 940–945 (1996).
- Sequist, E. R., Goetz, F. C., Rich, S. & Barbosa, J. Familial clustering of diabetic kidney disease. Evidence for genetic susceptibility to diabetic nephropathy. *N Engl J Med* **320**, 1161–1165 (1989).
- Wong, T. Y. *et al.* Diabetic retinopathy in a multi-ethnic cohort in the United States. *Am J Ophthalmol* **141**, 446–455, doi: 10.1016/j.ajo.2005.08.063 (2006).
- Pettitt, D. J., Saad, M. F., Bennett, P. H., Nelson, R. G. & Knowler, W. C. Familial predisposition to renal disease in two generations of Pima Indians with type 2 (non-insulin-dependent) diabetes mellitus. *Diabetologia* **33**, 438–443 (1990).
- Rema, M., Saravanan, G., Deepa, R. & Mohan, V. Familial clustering of diabetic retinopathy in South Indian Type 2 diabetic patients. *Diabet Med* **19**, 910–916, doi: 820 (2002).
- Hu, C. *et al.* CPVL/CHN2 genetic variant is associated with diabetic retinopathy in Chinese type 2 diabetic patients. *Diabetes* **60**, 3085–3089, doi: 10.2337/db11-0028 (2011).
- Hu, C. *et al.* PPARG, KCNJ11, CDKAL1, CDKN2A-CDKN2B, IDE-KIF11-HHEX, IGF2BP2 and SLC30A8 are associated with type 2 diabetes in a Chinese population. *PLoS one* **4**, e7643, doi: 10.1371/journal.pone.0007643 (2009).
- Hu, C. *et al.* Variations in KCNQ1 are associated with type 2 diabetes and beta cell function in a Chinese population. *Diabetologia* **52**, 1322–1325, doi: 10.1007/s00125-009-1335-6 (2009).
- Yoon, K. H. *et al.* Epidemic obesity and type 2 diabetes in Asia. *Lancet* **368**, 1681–1688, doi: 10.1016/S0140-6736(06)69703-1 (2006).
- Torrens, J. I. *et al.* Ethnic differences in insulin sensitivity and beta-cell function in premenopausal or early perimenopausal women without diabetes: the Study of Women's Health Across the Nation (SWAN). *Diabetes care* **27**, 354–361 (2004).
- Chan, J. C. *et al.* Diabetes in Asia: epidemiology, risk factors, and pathophysiology. *JAMA* **301**, 2129–2140, doi: 10.1001/jama.2009.726 (2009).
- McDonough, C. W. *et al.* A genome-wide association study for diabetic nephropathy genes in African Americans. *Kidney Int* **79**, 563–572, doi: 10.1038/ki.2010.467 (2011).
- Sheu, W. H. *et al.* Genome-wide association study in a Chinese population with diabetic retinopathy. *Hum Mol Genet* **22**, 3165–3173, doi: 10.1093/hmg/ddt161 (2013).
- Hanson, R. L. *et al.* Identification of PVT1 as a candidate gene for end-stage renal disease in type 2 diabetes using a pooling-based genome-wide single nucleotide polymorphism association study. *Diabetes* **56**, 975–983, doi: 10.2337/db06-1072 (2007).
- Shimazaki, A. *et al.* Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. *Diabetes* **54**, 1171–1178, doi: 54/4/1171 (2005).
- Huang, Y. C. *et al.* Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology* **118**, 642–648, doi: 10.1016/j.ophtha.2010.07.020 (2011).
- Fu, Y. P. *et al.* Identification of Diabetic Retinopathy Genes through a Genome-Wide Association Study among Mexican-Americans from Starr County, Texas. *J Ophthalmol* **2010**, doi: 10.1155/2010/861291 (2010).
- Palmer, N. D. *et al.* Evaluation of candidate nephropathy susceptibility genes in a genome-wide association study of African American diabetic kidney disease. *PLoS one* **9**, e88273, doi: 10.1371/journal.pone.0088273 (2014).
- Zhang, X. *et al.* Prevalence of diabetic retinopathy in the United States, 2005–2008. *JAMA* **304**, 649–656, doi: 10.1001/jama.2010.1111 (2010).
- Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678, doi: 10.1038/nature05911 (2007).
- Broadbent, H. M. *et al.* Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet* **17**, 806–814, doi: 10.1093/hmg/ddm352 (2008).
- Cugino, D. *et al.* Type 2 diabetes and polymorphisms on chromosome 9p21: a meta-analysis. *Nutr Metab Cardiovasc Dis* **22**, 619–625, doi: 10.1016/j.numecd.2010.11.010 S0939-4753(10)00294-2 (2012).
- Emanuele, E. *et al.* Chromosome 9p21.3 genotype is associated with vascular dementia and Alzheimer's disease. *Neurobiol Aging* **32**, 1231–1235, doi: 10.1016/j.neurobiolaging.2009.07.003 S0197-4580(09)00233-4 (2011).



33. Burdon, K. P. *et al.* Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. *Nat Genet* **43**, 574–578, doi: 10.1038/ng.824 (2011).
34. Schaefer, A. S. *et al.* Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. *PLoS Genet* **5**, e1000378, doi: 10.1371/journal.pgen.1000378 (2009).
35. Duesing, K. *et al.* Strong association of common variants in the CDKN2A/CDKN2B region with type 2 diabetes in French Europeans. *Diabetologia* **51**, 821–826, doi: 10.1007/s00125-008-0973-4 (2008).
36. Takeuchi, F. *et al.* Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* **58**, 1690–1699, doi: 10.2337/db08-1494 (2009).
37. Wu, Y. *et al.* Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes* **57**, 2834–2842, doi: 10.2337/db08-0047 (2008).
38. Cunningham, M. S., Santibanez Koref, M., Mayosi, B. M., Burn, J. & Keavney, B. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. *PLoS Genet* **6**, e1000899, doi: 10.1371/journal.pgen.1000899 (2010).
39. Yap, K. L. *et al.* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**, 662–674, doi: 10.1016/j.molcel.2010.03.021 (2010).
40. Zhao, Q. *et al.* Cross-sectional and longitudinal replication analyses of genome-wide association loci of type 2 diabetes in Han Chinese. *PLoS one* **9**, e91790, doi: 10.1371/journal.pone.0091790 (2014).
41. Lewis, J. P. *et al.* Association analysis in african americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. *Diabetes* **57**, 2220–2225, doi: 10.2337/db07-1319 (2008).
42. Horikoshi, M. *et al.* Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia* **50**, 2461–2466, doi: 10.1007/s00125-007-0827-5 (2007).
43. Steinhorsdottir, V. *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* **39**, 770–775, doi: ng2043 10.1038/ng2043 (2007).
44. Wilkinson, C. P. *et al.* Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology* **110**, 1677–1682, doi: S0161-6420(03)00475-5 (2003).
45. Wang, K. *et al.* A genome-wide association study on obesity and obesity-related traits. *PLoS one* **6**, e18939, doi: 10.1371/journal.pone.0018939 (2011).
46. Grant, S. F. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* **38**, 320–323, doi: 10.1038/ng1732 (2006).
47. Voight, B. F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* **42**, 579–589, doi: 10.1038/ng.609 (2010).
48. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265, doi: 10.1093/bioinformatics/bth457 (2005).
49. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575, doi: 10.1086/519795 (2007).

## Acknowledgments

This study was supported by Chinese National Natural Science Foundation of China grants 81070576 (W.D.L.), 81072922 (L.C.) and by grant 12JCZDJC24700 from Tianjin Municipal Science and Technology Commission (W.D.L.). We thank all of the patients for their cooperation and also the clinical doctors of Metabolic Diseases Hospital.

## Author Contributions

W.D.L., L.C. and J.L. conceived and designed the study; W.D.L., S.L. and F.W. wrote the manuscript; F.W., C.C., J.L., W.S., C.L., P.Y., H.J., F.Y., M.L., Y.W., L.Z., J.S. and M.C. performed the experiments; F.W., C.C., J.L., H.J., F.Y., W.S., P.Y., C.L. and W.D.L. analyzed the data; S.F., B.C., H.Z., H.H., Y.C., N.S., J.T., L.S., J.L., R.L., H.Q., S.L., J.L. and L.C. collected subjects and clinical data.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Wei, F. *et al.* TOX and CDKN2A/B Gene Polymorphisms Are Associated with Type 2 Diabetes in Han Chinese. *Sci. Rep.* **5**, 11900; doi: 10.1038/srep11900 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>