Familial Young-Onset Diabetes, Pre-Diabetes and Cardiovascular Disease Are Associated with Genetic Variants of DACH1 in Chinese

Ronald Ching Wan Ma1,2,3*, Heung Man Lee1,3,4, Vincent Kwok Lim Lam1,3,4, Claudia Ha Ting Tam1, Janice Siu Ka Ho1, Hai-Lu Zhao1, Jing Guan1, Alice Pik Shan Kong1,2,3, Eric Lau1, Guozhi Zhang1, Andrea Luk1, Ying Wang1, Stephen Kwok Wing Tsui4, Ting Fung Chan5, Cheng Hu6, Wei Ping Jia6, Kyong Soo Park7, Hong Kyu Lee8a, Hirotu Furuta6, Kishio Nanjo8, E. Shyong Tai9, Daniel Peng-Keat Ng9, Nelson Leung Sang Tang10,3, Jean Woo1, Ping Chung Leung11,9, Jeffrey Wong12, Po Sing Leung4, Terrence C. K. Lau*nb, Peter Chun Yip Tong1,2, Gang Xu1,2,3, Maggie Chor Yin Ng1c, Wing Yee So1,2,3, Juliana Chung Ngor Chan1,2,3*

1Department of Medicine and Therapeutics, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 2Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Hong Kong SAR, People's Republic of China, 3Li Ka Shing Institute of Health Science, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 4School of Biomedical Science, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 5School of Life Science, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 6Department of Endocrinology and Metabolism, Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Shanghai Clinical Center for Diabetes, Shanghai Kai Key Clinical Center for Metabolic Disease, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, People's Republic of China, 7Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology and Department of Internal Medicine, College of Medicine, Seoul National University, Chongno-gu, Seoul, Korea, 8First Department of Medicine, Wakayama Medical University, Wakayama, Japan, 9Department of Epidemiology and Public Health, National University of Singapore, Singapore, 10Department of Chemical Pathology, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 11Department of Orthopaedics and Traumatology, The Chinese University of Hong Kong, The Prince of Wales Hospital, Shatin, Hong Kong SAR, People's Republic of China, 12Department of Biochemistry, Hong Kong University of Science and Technology, Hong Kong SAR, People's Republic of China

Abstract

In Asia, young-onset type 2 diabetes (YOD) is characterized by obesity and increased risk for cardiovascular disease (CVD). In a genome-wide association study (GWAS) of 99 Chinese obese subjects with familial YOD diagnosed before 40-year-old and 101 controls, the T allele of rs1408888 in intron 1 of DACH1(Dachshund homolog 1) was associated with an odds ratio (OR) of 2.49(95% confidence intervals:1.57–3.96, P = 8.4×10−5). Amongst these subjects, we found reduced expression of DACH1 in peripheral blood mononuclear cells (PBMC) from 63 cases compared to 65 controls (P = 0.02). In a random cohort of 1468 cases and 1485 controls, amongst top 19 SNPs from GWAS, rs1408888 was associated with type 2 diabetes with a global P value of 0.0176 and confirmation in a multiethnic Asian case-control cohort (7370/7802) with an OR of 1.07(1.02–1.12, Pmeta = 0.012). In 599 Chinese non-diabetic subjects, rs1408888 was linearly associated with systolic blood pressure and insulin resistance. In a case-control cohort (n = 953/953), rs1408888 was associated with an OR of 1.54(1.07–2.22, P = 0.019) for CVD in type 2 diabetes. In an autopsy series of 173 non-diabetic cases, TT genotype of rs1408888 was associated with an OR of 3.31(1.19–9.19, P = 0.0214) and 3.27(1.25–11.07, P = 0.0184) for coronary heart disease (CHD) and coronary arteriosclerosis. Bioinformatics analysis revealed that rs1408888 lies within regulatory elements of DACH1 implicated in islet development and insulin secretion. The T allele of rs1408888 of DACH1 was associated with YOD, prediabetes and CVD in Chinese.

Citation: Ma RCW, Lee HM, Lam VKL, Tam CHT, Ho JSK, et al. (2014) Familial Young-Onset Diabetes, Pre-Diabetes and Cardiovascular Disease Are Associated with Genetic Variants of DACH1 in Chinese. PLoS ONE 9(1): e84770. doi:10.1371/journal.pone.0084770

Editor: Qingyang Huang, Central China Normal University, China

Received: July 3, 2013; Accepted: November 19, 2013; Published: January 20, 2014

Copyright: © 2014 Ma 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 partly supported by the Research Grant Council (RGC) Central Allocation Scheme (CUHK 1/04C), RGC earmarked grant (CUHK4462/06M), US National Institutes of Health (U01-RO15545-05) as well as the Hong Kong Foundation for Research and Development in Diabetes and Liao Wun Yuk Diabetes Research Memorial Fund established under the CUHK. The research in Shanghai Chinese was supported by grants from National Natural Science Foundation of China (30633001) and Shanghai Key Laboratory of Diabetes Mellitus (08DZZ223000). The study in Korea was funded by a grant from the Korea Health 21 R&D Project, Ministry of Health, Welfare and Family Affair, Republic of Korea (00-PJ3-PG6-GN07-001 to K.S.P.). The work in Japan was supported by Grant-in-Aid for Scientific Research on Priority Areas ‘Applied Genomics’ no. 17019047 from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The Singapore Prospective Study Program was funded through grants from the Biomedical Research Council of Singapore (BMRC 05/136/19/413 and 03/1/27/18/216) and the National Medical Research Council of Singapore (NMRC/N01/C07/14/2008). The Singapore Malay Eye Study was funded by the National Medical Research Council (NMRC 0796/2003, IRRG07nov013, and NMRC/SStaR/ 0003/2008) and Biomedical Research Council (BMRC, 09/1/35/19/616). EST also receives additional support from the National Medical Research Council through a clinician scientist award. The Singapore BioBank and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services for tissue archival and genotyping, respectively. 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: jchan@cuhk.edu.hk
† These authors contributed equally to this work.
‡ Current address: Department of Internal Medicine, Eulji University, Hangeulbiiseok-gil, Hayge-dong, Nowon-gu, Seoul, Korea
§ Current address: Department of Biology and Chemistry, City University of Hong Kong, Hong Kong SAR, People’s Republic of China
¶ Current address: Center for Diabetes Research, Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America.
Introduction

Genome-wide association studies (GWAS) and their meta-analyses have discovered novel loci for T2D in European [1] and Asian populations with odds ratios (ORs) of 1.1–1.4 [2,3,4]. In Asia, the most rapid increase in diabetes occurs in the young-to-middle aged group. Using Hong Kong Chinese as an example, 20% of people with diabetes were diagnosed before the age of 40 years. In these young patients, less than 10% had type 1 presentation and 15% had monogenic diabetes. In the remaining patients, family history, obesity and premature cardiovascular disease (CVD) were prominent features [5,6,7]. In the Hong Kong Family Diabetes Study (HKFS) which recruited family members of patients with young-onset diabetes (YOD), we reported strong heritability of diabetes and obesity [8] with co-linkage of related traits to multiple chromosomal regions including chromosome 1q [9]. However, the genetic basis of this form of YOD has not been studied.

Epidemiological analysis has confirmed the clustering of metabolic syndrome, insulin resistance, diabetes and CVD [10,11], which may share common genetic, environmental or lifestyle factors [1]. In a meta-analysis of 3 GWAS conducted in Chinese from Hong Kong and Shanghai, we discovered rs10229583 located in 7q32 near PAX4 which was subsequently confirmed in Asian and Caucasian populations [12]. In the first discovery cohort of this meta-analysis consisting of 99 Hong Kong Chinese patients with YOD diagnosed before 40-year-old with at least 1 affected first degree relative and obesity, the T allele of rs1408886 in intron 1 of DACH1 (Dachshund homolog 1) with $P < 10^{-5}$ was replicated in a multiethnic Asian case-control cohort, albeit insignificant in Caucasians.

In Drosophila, Dachshund (DAC) is the homolog of DACH1 in human which is a highly conserved transcription factor implicated in developmental biology [13,14]. In a recent report, DAC was found to interact physically with PAX6 to control insulin expression [14]. In light of these findings, we revisited the risk association of rs1408886 of DACH1 and explored whether this genetic variant might be associated with YOD and related traits in Chinese populations. To test this hypothesis, we examined the differential expression of DACH1 in peripheral blood mononuclear cells (PBMC) in patients with YOD and tested the genetic associations of rs1408886 in multiple case-control cohorts followed by bioinformatic analysis. We found reduced expression of DACH1 in PBMC from subjects with YOD and association of the T allele of rs1408886 with YOD, prediabetes and CVD. This risk variant is located within the vicinity of conserved non-coding elements (CNE) of DACH1 associated with multiple consensus transcription factor binding sites and chromatin modification sites suggesting possible regulatory functions. These findings suggested that genetic variants of DACH1 may be implicated in abnormal islet biology resulting in prediabetes, YOD and CVD in Chinese populations.

Results

Associations with familial young-onset T2D in GWAS

In the GWAS of 99 obese Chinese subjects with familial YOD and 101 controls, 423,513 of 541,891 autosomal SNPs passed quality control with no population stratification using multidimensional scaling analysis and after adjusting for genomic control (GC). From stage 1 GWAS, 24 unique loci with the lowest $P$ value in the dataset ($P < 10^{-5}$) were taken forward for replication in 1468 cases and 1485 controls. Of these, 19 SNPs passed the quality control criteria and 2 SNPs (rs1408888 and rs1449675) remained significantly associated with T2D (Table 1). The intronic SNP rs1408888 {stage 1: OR [95% confidence intervals (CI)] = 2.49(1.57–3.96), $P = 8.4 \times 10^{-5}$; stage 2: 1.15(1.03–1.29), $P = 0.0164$} was located at chromosome 13q21.3 and lies within the first intron of the DACH1 gene (Figure 1). The intergenic SNP rs1449675 {stage 1: OR = 5.33(2.30–12.36); $P = 2.0 \times 10^{-5}$, 1.19(1.00–1.41); $P = 0.0439$} was located at chromosome 6q25.3. In the combined analysis (1567 cases, 1566 controls), three more SNPs (rs6095551 in ZNF608, rs987105 in MUT, and rs1413119 in an intergenic region on chromosome 13) showed nominal associations (OR < 0.05). Among these five SNPs, rs1408886 in DACH1 had the highest OR [1.21(1.08–1.35); $P = 9.1 \times 10^{-4}$] which remained significant ($P = 0.0176$) after correction for multiple testing of the 19 SNPs using 10,000 permutations. In a meta-analysis of 5 Asian case-control cohorts consisting of 7370 cases and 7802 controls, we obtained a combined OR of 1.07(1.02–1.12, $P = 0.0112$) with no heterogeneity ($I^2 = 0.107$) in Cochran’s Q test and $I^2 = 44.8\%$ (0.0%–76.4%) (Table 2).

Associations with quantitative traits in healthy adults

In 599 healthy adults and after adjustment for age and gender, systolic blood pressure (BP), Homeostasis Model Assessment index for insulin resistance (HOMA-IR) and $\beta$ cell function (HOMA-$\beta$) and fasting insulin were associated with increasing number of alleles. Using multivariate analysis, the T-allele of DACH1 was associated with systolic BP [1.56(1.02–2.10) per T-allele, 0.61% variance explained], fasting plasma insulin [1.072(0.006–0.151) per T-allele, 1.03% variance explained] and HOMA-IR [1.067 (0.012–0.145) per T-allele, 0.97% variance explained] (Table 3).

Clinical and pathological association with cardiovascular disease (CVD)

In the matched case-control cohort, Chinese T2D patients with CVD were more obese and had worse dyslipidaemia and renal function than those without CVD (Table 4). The TT/TG genotype was associated with an OR of 1.51 ($P = 0.02$) which remained significant after adjusting for estimated glomerular filtration rate (eGFR) with an OR of 1.54 (1.07–2.22, $P = 0.019$). In an autopsy series of 173 non-diabetic cases, rs1408886 did not depart from Hardy-Weinberg Equilibrium (HWE). Compared to cases with TG/GG genotype (n = 83, mean ± SD age: 70.6 ± 15.7 years, 43% female), cases with TT genotype (n = 90, age: 67.0 ± 15.7 years, 38% female) were more likely to have a history of coronary heart disease (CHD) (17.8% versus 7.2%, $P = 0.0375$) and coronary arteriosclerosis (16.7% versus 6.0%, $P = 0.0287$). The respective ORs were 3.31 (1.19–9.19, $P = 0.0214$) and 5.27 (1.25–11.07, $P = 0.0184$) after adjustment for age and sex.

Bioinformatics analysis and expression study

Two neighboring SNPs in weak linkage disequilibrium (LD) ($r^2 = 0.5$) with rs1408888 (rs9572813 and rs17791181) also showed nominal association with T2D ($P = 0.01–0.001$) in the GWAS analysis (Figure 1). Bioinformatics analysis revealed that the region between rs1408888 and rs9572813 overlapped with a regulatory element conserved from fugu fish to human [15]. On datamining, this element [OREG0002711 (http://www.oreganno.org/oregano/) or chr13:72,425,787-72,428,335 (hg19) (http://enhancer.lbl.gov/frnt_page_n.shtml)] shows an enhancer activity which directs the distinct expression of a β-galactosidase reporter gene in the eye, cranial nerve, forebrain, hindbrain and neural tube in the mouse embryos [15,16]. In this region, another non-
| SNP          | Chr. | Nearest gene(s) | Risk allele | Stage | RAF (T2D) | RAF (Controls) | HapMap CHB RAF | OR (95% CI)     | Combined P Allelic | OR (95% CI)     | Ppermutation |
|-------------|------|-----------------|-------------|-------|-----------|---------------|---------------|----------------|-------------------|-----------------|--------------|
| rs841859    | 1    | SLC2A1          | G           | 1     | 0.237     | 0.089         | 0.173         | 3.18 (1.77 - 5.71) | 5.8 x 10^{-5}   | 1.12 (0.97 - 1.29) | 0.1386       | 0.9422      |
| rs6661853   | 1    | CNH3            | G           | 1     | 0.798     | 0.614         | 0.673         | 2.48 (1.59 - 3.89) | 5.4 x 10^{-5}   | 0.97 (0.87 - 1.09) | 0.6128       | 1.0000      |
| rs16862964  | 3    | LPP             | G           | 1     | 0.480     | 0.262         | 0.323         | 2.59 (1.71 - 3.94) | 6.7 x 10^{-6}   | 0.97 (0.88 - 1.08) | 0.5845       | 1.0000      |
| rs4834621   | 4    | G               | 1           | 2     | 0.216     | 0.233         | 0.91 (0.8 - 1.03) | 0.1238          |                   |                |              |
| rs7665789   | 4    | A               | 1           | 2     | 0.829     | 0.842         | 0.91 (0.79 - 1.05) | 0.1975          |                   |                |              |
| rs6595551   | 5    | ZNF608          | G           | 1     | 0.748     | 0.530         | 0.584         | 2.63 (1.72 - 4.01) | 5.9 x 10^{-6}   | 1.12 (1.01 - 1.24) | 0.0337       | 0.4836      |
| rs130932    | 6    | POU5F1          | C           | 1     | 0.460     | 0.272         | 0.387         | 2.27 (1.5 - 3.45)  | 1.0 x 10^{-4}   | 1.05 (0.94 - 1.16) | 0.3990       | 0.9999      |
| rs84514     | 6    | LRFN2           | A           | 1     | 0.849     | 0.678         | 0.679         | 2.66 (1.63 - 4.33) | 6.3 x 10^{-5}   | 1.04 (0.93 - 1.17) | 0.4618       | 0.9999      |
| rs987105    | 6    | MUT             | G           | 1     | 0.939     | 0.807         | 0.911         | 3.71 (1.88 - 7.32) | 7.1 x 10^{-5}   | 1.25 (1.06 - 1.47) | 0.0075       | 0.1388      |
| rs1325076   | 6    | FUT9            | G           | 1     | 0.444     | 0.243         | 0.232         | 2.5 (1.63 - 3.83)  | 2.1 x 10^{-5}   | 1.02 (0.92 - 1.14) | 0.7021       | 1.0000      |
| rs1449675   | 6    | CTNNA3          | A           | 1     | 0.965     | 0.837         | 0.893         | 5.33 (2.3 - 12.36) | 2.0 x 10^{-5}   | 1.29 (1.09 - 1.52) | 0.0025       | 0.0503      |
| rs10762033  | 10   | CTNNA3          | G           | 1     | 0.566     | 0.366         | 0.476         | 2.25 (1.51 - 3.36) | 6.4 x 10^{-5}   | 1.07 (0.97 - 1.18) | 0.2078       | 0.9897      |
| rs4245124   | 11   | SPATA19         | C           | 1     | 0.697     | 0.485         | 0.565         | 2.44 (1.62 - 3.68) | 1.7 x 10^{-5}   | 1.07 (0.96 - 1.18) | 0.2213       | 0.9922      |
| rs1413119   | 13   | DOCK1           | A           | 1     | 0.818     | 0.644         | 0.768         | 2.49 (1.57 - 3.96) | 8.4 x 10^{-5}   | 1.21 (1.08 - 1.35) | 9.1 x 10^{-4} | 0.0176      |
| rs11069344  | 13   | DOCK1           | G           | 1     | 0.318     | 0.139         | 0.244         | 2.9 (1.76 - 4.78)  | 1.8 x 10^{-5}   | 1.05 (0.92 - 1.18) | 0.4860       | 0.9999      |
| rs1165227   | 17   | MS2             | G           | 1     | 0.636     | 0.436         | 0.577         | 2.27 (1.52 - 3.39) | 5.7 x 10^{-5}   | 0.98 (0.88 - 1.08) | 0.6715       | 1.0000      |
Joint analysis of stage 1 + stage 2

| Stage 1 | Stage 2 | Combined |
|---------|---------|----------|
| P       | OR      | P        |
| 0.546   | 0.566   | 0.92 (0.83 – 1.02) | 0.1328 |
| 1.01 (0.92 – 1.12) | 0.7849 |
| 0.626   | 0.421   | 0.47 (2.31 (1.54 – 3.41) | 3.9 |
| 0.518   | 0.528   | 0.96 (0.87 – 1.06) | 0.4301 |
| 1.09 (0.98 – 1.21) | 0.1177 |
| 0.717   | 0.525   | 0.702 (2.3 (1.52 – 3.48) | 7.4 |
| 0.646   | 0.639   | 1.03 (0.93 – 1.15) | 0.5715 |

**Nearest Entrez genes within 250 Kbp.**

**Discussion**

In an adequately powered case-control cohort of familial YOD characterized by obesity, we discovered several common SNPs (allele frequency > 0.3) with ORs of 2–2.5, including the T allele of rs1408888 of DACH1 with allele frequency of 0.75 and an OR of 2.49. In a subset of this cohort, we found reduced expression of DACH1 in PBMC patients with YOD. In the stage 2 experiment, we successfully genotyped 19 SNPs with the lowest P value and replicated the association of rs1408888 with an OR of 1.21 and a global P < 0.05 in a random case-control cohort of 2953 Chinese with older age of diagnosis. This was followed by confirmation in 15,172 Asian subjects with an OR of 1.07. We further demonstrated that the T allele was associated with insulin resistance and high BP in normal subjects, CVD in patients with T2D, and coronary arteriosclerosis in autopsy samples. Bioinformatics analysis revealed its location in a conserved region within the intron of DACH1, implicated in pancreatic islet development [13]. Unlike the novel SNP rs10229583 at 7q32 near PAX4 discovered in a meta-analysis of 3 GWAS comprising 684 T2D patients and 955 controls of Southern Han Chinese descent with confirmation in a multi-ethnic population (Pmeta = 2.3 × 10^-10 in East Asians; P = 8.6 × 10^-7 in Caucasians) [12], the risk association of rs1408888 with T2D was increasingly attenuated in a multi-ethnic population, suggesting that this SNP might be more relevant to Asians undergoing rapid transition characterized by obesity and young age of diagnosis [6].

**Known function of DACH1**

DACH1, located on chromosome 13q21, is the mammalian homologue of the *Drosophila* dachshund (Dac) gene. It encodes a well-conserved nuclear protein capable of binding to DNA and has two highly conserved domains (DachBox-N and DachBox-C) from *Drosophila* to humans [19, 20]. It is a key component of the retinal determination gene network that governs cell fate and plays a key role in ocular, limb, brain and gonadal development [21]. DACH1 knockout mice die shortly after birth, with no gross histological abnormalities observed in eyes, limbs, or brain tissues, suggesting its possible role in perinatal development [22]. Herein, DACH1 is a critical transcription factor in tissue differentiation and organ coding element (CNE803) [17], highly conserved in vertebrates, shows homology to an EST from the human eye (BY793700) (Figure 1). We sequenced the region between rs1408888 and rs9572813 in 74 controls and 82 cases who had GWAS data and did not discover novel variants in the CNE803 nor in the surrounding regions. In this genomic region containing multiple SNPs, 3 SNPs (rs17252745, rs17252752 and rs57143718) with marked inter-ethnic differences in the NCBI SNP database (Figure 1) (<http://www.ncbi.nlm.nih.gov/snp/>) [minor allele frequency (MAF) < 0.05 in CEU (Caucasians) versus 0.4–0.5 in CHB/JPT (Asians)] showed normal association with YOD with OR ranging from 1.36 to 1.53. They were further genotyped in 206 controls and 390 subjects with YOD with one of the SNPs (T allele of rs57143718) showing association in the combined cohort of 472 cases and 280 controls [OR = 1.26 (1.02–1.56), P = 0.036]. We used reverse transcription PCR and Northern blot analysis to examine the expression of CNE803 in PBMC, pancreatic progenitor cells (PPC) [18], HCT116 and HKCI-2 cancer cells which was negative. Expression of DACH1 was detected in PBMC and PPC. Using quantitative real-time PCR, lower DACH1 expression was found in PBMC from patients with YOD compared to control subjects (P = 0.02 for regression with age and sex adjustment, P = 0.005 for Wilcoxon rank sum test) (Figure 2).
Figure 1. **Upper panel:** Regional plot showing significant association of rs1408888 in the DACH1 locus. The $-\log_{10} P$ values for the allelic test from stage 1 (genome scan) were plotted as a function of genomic position (NCBI build 36). Rs1408888 which showed the strongest signal and neighboring genotyped SNPs in the joint analysis were denoted by purple diamond. LD information (based on HapMap) was shown by color-coded points. Two neighboring SNPs rs9572813 and rs17791181, which showed nominal significance and moderate linkage disequilibrium ($0.4 < r^2 < 0.6$) with rs1408888 were indicated. Estimated recombination rate (the blue line) based on the Japanese and Chinese HapMap population was plotted to reflect the local LD structure around the significant SNPs. Gene annotations were taken from NCBI. **Lower panel:** Bioinformatics analysis of genomic region surrounding rs1408888. The region harboring rs1408888 lies in close vicinity of 2 highly conserved non-coding elements, CNE803 and OREG0002711. The two blue arrowheads at the end indicate the positions of rs1408888 (right blue dot) and rs9572813 (left blue dot). The three internal red arrows indicate the positions of the three SNPs (rs17252745, rs17252752 and rs57143718, red dots from left to right) genotyped by sequencing. The allele frequencies of the SNPs are shown in pie chart at the bottom. The alignment of the highly conserved fugu CNE803, the human sequence corresponding to the fugu CNE and the eye prepared EST BY797940 are also shown.

doi:10.1371/journal.pone.0084770.g001
development. One of its gene targets is the mothers against decapentaplegic homolog 4 (SMAD4), which upon binding with DACH1 can result in repression of TGFβ signaling and TGFβ-induced apoptosis [23]. In this regard, increased TGFβ1 activity has been implicated in heart and vascular development, hypertension and progressive myocardial fibrosis [24,25]. In a recent proof-of-concept analysis, researchers merged co-expression and interaction networks and detected/inferred novel networks to which facilitated oxidative stress with expression of early transcription genes via MAPK pathway leading to cardiovascular and renal complications. The second network proposed the interaction of beta-catenin with CDH5 and TGFBR1 through Smad molecules to contribute to endothelial dysfunction [26], often a precursor of CVD [27].

**Table 2. Meta-analysis of risk association of DACH1 rs1408888 with Type 2 diabetes in independent multi-ethnic Asian case-control cohorts.**

| Study          | n   | T2D | Control | Total | Risk allele frequency |
|----------------|-----|-----|---------|-------|-----------------------|
| Hong Kong Chinese | 1567| 1586| 3153    | 0.753 | 21.9%                 |
| Shanghai Chinese  | 1779| 1833| 3612    | 0.763 | 24.57%                |
| Korean           | 749 | 616 | 1365    | 0.560 | 8.05%                 |
| Singapore Chinese | 2010| 1945| 3955    | 0.762 | 24.07%                |
| Singapore Malay  | 794 | 1240| 2034    | 0.673 | 12.94%                |
| Japanese         | 471 | 582 | 1053    | 0.666 | 8.46%                 |
| Asian meta-analysis | 7370| 7802| 15172   | --    | 1.07 (1.02 – 1.12)    |

**Table 3. Clinical and metabolic characteristics of Hong Kong Chinese healthy adults stratified according to the genotypes of DACH1 rs1408888.**

| Characteristics | GG (N = 55) | TG (N = 246) | TT (N = 298) | P   |
|-----------------|-------------|--------------|--------------|-----|
| Body mass index (kg/m²) | 22.7±3.8 | 22.9±3.3 | 23.0±3.3 | 0.801 |
| Waist circumference (cm) | 74.7±10.7 | 76.7±9.8 | 77.2±9.1 | 0.780 |
| Hip circumference (cm) | 92.8±6.4 | 93.3±6.3 | 93.6±5.8 | 0.624 |
| Systolic BP (mmHg) | 111±14.5 | 114.3±16.1 | 116.8±16.9 | 0.030 |
| Diastolic BP (mmHg) | 68.6±10.8 | 72.2±11 | 72.9±11.5 | 0.073 |
| Total cholesterol (mmol/l) | 4.7±0.8 | 5±0.9 | 5.1±1 | 0.025 |
| Triglyceride (mmol/l) | 0.7 (0.6 – 1.1) | 0.9 (0.6 – 1.3) | 0.9 (0.7 – 1.3) | 0.547 |
| HDL-C (mmol/l) | 1.5±0.4 | 1.6±0.4 | 1.5±0.4 | 0.306 |
| LDL-C (mmol/l) | 2.8±0.8 | 3±0.8 | 3±0.9 | 0.071 |
| Fasting plasma glucose (mmol/l) | 4.9 (4.5 – 5.2) | 4.8 (4.6 – 5.1) | 4.8 (4.6 – 5.1) | 0.709 |
| Fasting plasma insulin (pmol/l) | 39.9 (23.1 – 51.2) | 40.2 (24.4 – 55.9) | 42.4 (29.7 – 61.6) | 0.014 |
| HOMA-IR | 1.4 (0.8 – 1.9) | 1.5 (0.9 – 2.0) | 1.5 (1 – 2.2) | 0.019 |
| HOMA-β | 93.6 (68.1 – 135.6) | 99 (63.4 – 162.6) | 112.5 (71.0 – 167.6) | 0.010 |
| Insulinogenic index (mU/mmol) | 13.9 (8.1 – 21.4) | 15.5 (9.3 – 23.6) | 16.3 (9.7 – 25.6) | 0.2369 |

Data are expressed as n or mean±SD or median (interquartile range). P values were calculated from linear regression adjusted for sex and age assuming an additive model. Homeostasis model assessment of insulin resistance (HOMA-IR) was calculated as (FPI x FPG)/22.5, and homeostasis model assessment of beta-cell function (HOMA-β) was calculated as FPI x 20 / (FPG – 3.5) where FPI = fasting plasma insulin and FPG = fasting plasma glucose. Beta cell function (×10⁻⁹) was determined by the formula: [insulin AUCᵐ⁻³ (min:mmol)/ glucose AUCᵐ⁻³ (min:mmol)] where AUC = area under curve. Insulinogenic index was calculated as (FPI during OGTT for 30 min – 0 min) / (FPG during OGTT for 30 min – 0 min) where OGTT = 75 gram oral glucose tolerance test.

doi:10.1371/journal.pone.0084770.t003
findings highlight the possible role of DACH1 in regulating functions of multiple organs.

In support of these genetic associations, in a mouse model of diet-induced ß-cell dysfunction, islet DACH1 gene expression was reduced in prediabetic animals fed a high-fat diet [32]. In both zebrafish and mice, loss of DACH1 in mice did not affect the number of PPC, it blocked the perinatal burst of proliferation of differentiated β-cells 

### Table 4. Clinical and metabolic characteristics of a case-control cohort of Hong Kong Chinese Type 2 diabetic patients with or without cardiovascular disease (CVD) matched for age, sex and disease duration.

|                          | No CVD         | CVD            | P values |
|--------------------------|----------------|----------------|----------|
| N                        | 953            | 953            |          |
| Sex (% of male)          | 48.3%          | 48.3%          |          |
| Age (years)              | 64.1±10.3      | 64.1±10.3      |          |
| Age of diagnosis (years) | 54.3±12.7      | 54.4±12.7      |          |
| Diabetes duration (years)| 9.4±6.7        | 9.4±6.7        |          |
| Body mass index (kg/m²)  | 24.4±4.6       | 24.7±4.9       | 0.124    |
| Waist circumference (cm) | 85.2±11.5      | 86.3±12.8      | 0.063    |
| Systolic blood pressure (mmHg) | 140.0±22.8 | 143.3±21.9 | <0.001 |
| Diastolic blood pressure (mmHg) | 76.0±12.2   | 76.9±11.7     | 0.110    |
| Glycated hemoglobin (%)  | 7.4(6.4–8.6)   | 7.8(6.8–9.3)   | <0.001   |
| High density lipoprotein cholesterol (mmol/L) | 1.27(1.04–1.51) | 1.17(1.0–1.4) | 0.004 |
| Low density lipoprotein cholesterol (mmol/L) | 3.20(2.6–3.8)  | 3.30(2.7–4.0) | 0.050   |
| Triglyceride (mmol/L)    | 1.37(0.93–2.0) | 1.52(1.05–2.23)| 0.053    |
| Urinary albumin:creatinine ratio (mg/mmol) | 2.72(0.91–14.3) | 7.5(1.80–49.3) | <0.001 |
| Estimated GFR (ml/min/1.73 m²) | 97.9(77.5–116.8) | 89.1(65.7–109.1) | <0.001 |

| DACH1 rs1408888  | Model   | OR(95%CI); P values | OR(95%CI); P values |
|-------------------|---------|---------------------|---------------------|
| TT count (%)      | 522(54.8)| 524(55.0)           | Dominant            |
| TG count (%)      | 353(37.0)| 376(39.5)           | Recessive           |
| GG count (%)      | 78(8.2) | 53(5.5)              | Allelic             |

Types of CVD (number, %)

|                          | No CVD | CVD  | P values |
|--------------------------|--------|------|----------|
| Coronary heart disease   | -      | 541  | 56.8%    |
| Stroke                   | -      | 453  | 47.5%    |
| Peripheral vascular disease | -    | 223  | 23.4%    |
| 1 vascular bed           | -      | 712  | 74.7%    |
| 2 vascular beds          | -      | 218  | 22.9%    |
| 3 vascular beds          | -      | 23   | 2.4%     |

Data are expressed in mean±SD or median(interquartile range) or n, %. *P values and ORs were estimated by the logistic regression with adjustment for logarithm of eGFR.

Cardiovascular diseases was diagnosed based on clinical history and assessment at enrolment to the Hong Kong Diabetes Registry and/or subsequent events defined by the International Classification of Diseases, Ninth Revision (ICD-9), retrieved from the Hong Kong Death Registry and Hong Kong Hospital Authority (HA) Central Computer System. Coronary heart disease (CHD) was defined as myocardial infarction (ICD-9 code 410), ischemic heart disease (ICD-9 code 411-414) or death due to CHD (ICD-9 code 410-414). Stroke was defined as non-fatal (ICD-9 code 432-434, 436) or fatal ischemic stroke (ICD-9 code 432-438), or, hemorrhagic stroke as defined by fatal and non-fatal subarachnoid hemorrhage (ICD-9 code 430), intracerebral hemorrhage (ICD-9 code 431) or other/unspecified intracranial hemorrhage (ICD-9 code 432). Peripheral vascular disease (PVD) was defined as ankle-brachial ratio <0.9 using Doppler ultrasound scan, diabetes with peripheral circulation disorders (ICD-9 code 250.7), gangrene (ICD-9 code 785.4), angiopathy in diseases classified elsewhere (ICD-9 code 443.81), peripheral vascular disease unspecified (ICD-9 code 443.9), other peripheral vascular shunt or bypass (procedure code 39.29), insertion of non-drug-eluting peripheral vessel stents (procedure code 39.90) or amputation of lower limb (procedure code 84.1) without a traumatic amputation diagnosis code (ICD-9 code 895-897).

PLOS ONE | www.plosone.org 7 January 2014 | Volume 9 | Issue 1 | e84770
DOI: 10.1371/journal.pone.0084770.t004

Given the strong links between T2D and CVD, the association of TT/TG genotype of DACH1 with CVD with an odds ratio of 1.54, after adjustment for age, sex, disease duration and eGFR was noteworthy. Patients with CVD were more obese and had more risk factors including high BP. Interestingly, in normal subjects, the T allele was linearly associated with BP and insulin levels which are well known risk factors for CVD [11]. In the autopsy series, TT carriers had 2–3 fold increased risk of coronary arteriosclerosis and CHD. These consistent findings in independent cohorts at different stages of the spectrum of cardio-metabolic disease, together with experimental studies from other groups, strongly support the role of DACH1 in these complex diseases.
The complexity of human evolution and natural selection by external forces, including but not limited to temperature, foods, infections, can result in diversity of genomic architecture and expression, making replication of genetic association of complex diseases challenging [33]. Given the rapid westernization of Hong Kong Chinese within less than a century, we hypothesize that DACH1 may be a thrifty gene which regulates growth to improve survival chances during time of hardship but increases risk of obesity, prediabetes, YOD and CVD during time of affluence [34]. Hitherto, apart from maturity onset diabetes of the young [35], the genetics of familial YOD characterized by obesity have not been well studied and these results might motivate further research in subjects with these phenotypes to confirm or refute our hypothesis.

Possible significance of rs1408888 of DACH1

The risk allele rs1408888 is located in the first intron of DACH1 within the vicinity of conserved elements [16], which can direct a unique gene expression pattern resembling the embryonic expression pattern of DACH1 [22]. One of these elements, CNE803 located 1.6 Kb from rs1408888 (Figure 1), showed sequence homology to an EST from an eye library (BY797940). We sequenced this region in the original GWAS cohort but did not find any novel SNPs. Three SNPs in this region which were sequence homology to an EST from an eye library (BY797940).

rs57143718 to the expression level of DACH1 may be a thrifty gene which regulates growth to improve survival chances during time of hardship but increases risk of obesity, prediabetes, YOD and CVD during time of affluence [34]. Hitherto, apart from maturity onset diabetes of the young [35], the genetics of familial YOD characterized by obesity have not been well studied and these results might motivate further research in subjects with these phenotypes to confirm or refute our hypothesis.

Figure 2. Expression of DACH1 detected by quantitative real-time PCR, in peripheral blood mononuclear cells (PBMC) extracted from 65 control subjects and 63 young-onset type 2 diabetic patients (DM). Expression level was normalized to the expression of β actin using the ΔΔCt method. The results are represented as mean ± standard error of the mean (SEM) with age and sex adjustment.

doi:10.1371/journal.pone.0084770.g002

Table 5. Clinical profiles of discovery, replication and validation cohorts in a 3-stage genome wide association study in Asian populations.

| Country        | Discovery cohort | Replication cohort | Validation cohort |
|----------------|------------------|--------------------|-------------------|
| Singapore Malay | T2D Controls     | T2D Controls       | T2D Controls      |
| Age (years)    | 40.6 ± 8.8       | 37.4 ± 6.8         | 38.3 ± 6.7        |
| Body mass index (kg/m²) | 24.8 ± 3.9       | 25.3 ± 3.5         | 25.5 ± 3.7        |
| HbA1C (%)      | 8.0 ± 3.6        | 8.0 ± 3.5          | 8.1 ± 3.1         |
| doi:10.1371/journal.pone.0084770.t005
importance of DACH1 and agrees with the bioinformatic analysis of the region surrounding rs1408888. The functional significance of this SNP/locus requires further exploration.

Limitations and conclusion
In this multi-staged experiment, we have discovered risk association of an intronic SNP (rs1408888) of DACH1 with YOD, BP, insulin resistance and CVD in Chinese populations. Although no significant association between rs1408888 and the YOD subgroup in the stage 1 replication cohort was found ($P > 0.1$), possibly due to small sample size, the age of diagnosis was relatively young in our discovery (31.8±7.7 years) and replication cohorts for T2D (44.0±13.6 years) and that for CVD (54.3±12.7 years) (Tables 4 and 5) compared to most GWAS. Given that these findings were mainly found in Asian population where YOD is a predominant feature, our findings highlight the need for more genetic research in YOD. Together with the known function of DACH1 on developmental biology and regulation of insulin secretion and its reduced expression in human PBMC associated with YOD, these findings add to the growing body of knowledge regarding the candidacy role of DACH1 for cardio-metabolic dysfunction, which may manifest as YOD in populations undergoing rapid transition in nutrition and lifestyles.

Methods
Risk association with diabetes using multiple cohorts
The clinical characteristics of the discovery and replication cohorts of the Asian population as well as methods of genotyping and genetic analysis have been reported [12]. In brief, the discovery cohort (stage 1) consisted of 99 obese Chinese subjects (BMI≥27 kg/m² or waist≥90 cm in men or ≥80 cm in women) with YOD (age of diagnosis<40 years) and at least one affected...
first-degree relative selected from the HKFDS [8,9] and the Hong Kong Diabetes Registry (HKDR) [36]. The 101 age- and sex-matched control subjects were selected from a community-based health promotion program [37]. The replication cohort (stage 2) consisted of 1468 T2D subjects selected from the HKDR [36] and unrelated subjects from the HKFDS [8] while the control cohort consisted of 507 healthy volunteers [37] and 978 adolescents [38].

The validation cohorts (stage 3) consisted of 1892 cases and 1808 controls from Shanghai [39]; 749 cases and 616 controls from Korea [40]; 2804 cases (2010 Chinese, 794 Malay) and 2185 controls (1945 Chinese, 1240 Malays) from Singapore [41] and 471 cases and 582 controls from Japan [40] (Table 5). Written informed consent was obtained from all participants or their parents with approval by the Clinical Research Ethics Committee of the Chinese University of Hong Kong, the ethics committee of the Wakayama Medical University, the institutional review boards of the Clinical Research Institute in the Korea Seoul National University Hospital and the Shanghai Jiao Tong University Affiliated Sixth People’s Hospital.

Quantitative traits, CVD and clinico-pathological features

All subjects in the Hong Kong control cohort had documentation of anthropometric indexes, BP, cardiovascular risk factors, plasma insulin and glucose during 75 gram oral glucose tolerance test [42]. The case-control cohort was selected from the HKDR set up in 1995 as part of a quality improvement program using structured protocols [36]. Using this cohort, we selected a case-control cohort of CVD (953/953) matched for age, sex and disease duration. Definitions of CVD (including ischaemic heart disease, stroke and peripheral vascular diseases) were based on clinical assessments and the International Classification of Disease 9th version. In an autopsy series with documentation of pathological features and clinical history [43,44], we extracted genomic DNA from archived paraffin blocks using white blood cell-concentrated spleen tissues for genotyping. Figure 3 summarizes the selection criteria and study flow.

Genotyping

Discovery cohort. The discovery cohort (99 cases and 101 controls) were assayed with Illumina HumanHap550-Duo BeadChip at deCODE Genetics. Of the 541,891 genotyped autosomal SNPs, 116,378 (21%) SNPs were excluded due to call rate <0.95 (n = 2311); MAF<0.05 (n = 113,596) and significant departure from HWE in control subjects (P<0.001) (n = 947), giving 425,513 SNPs on chromosome 1–22 for final analysis after checking for population stratification.

Replication cohort. In the second stage, 24 SNPs with the lowest P value (P<1×10−6 in allelic test) were genotyped in the replication cohort. Only one SNP was genotyped for locus with multiple SNPs in high LD (r²>0.6). Genotyping was performed at the McGill University and Genome Quebec Innovation Centre using primer extension of multiplex products with detection by MALDI-TOF mass spectroscopy on a Sequenom MassARRAY platform (San Diego, CA, USA). Of 24 genotyped SNPs, 5 were excluded due to low call rate (<90%). All remaining 19 SNPs were in HWE in controls with a concordance rate of 96% in 65 blinded duplicate samples.

Validation cohorts. rs1408888 which showed significance in stage 1, 2 and combined cohorts were genotyped in the Asian populations using the following methods: 1) Shanghai Chinese: MassARRAY platform (MassARRAY Compact Analyzer, Sequenom, San Diego, CA, USA) with 97.5% call rate and 100% concordance rate; 2) Korea: Assay-on-Demand TaqMan assays (Applied Biosystems, Foster City, CA, USA) and ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA) with 99.4% call rate and 100% concordance rate based on 13 duplicates; 3) Japan: TaqMan SNP genotyping system (Applied Biosystems, Foster City, CA) and ABI PRISM 7700 system with 20% of samples directly sequenced using Sanger sequencing and analyzed with an ABI 3100 capillary sequencer with 100% concordance rate. For the Singapore study, the Chinese samples were genotyped on the 610Quad and iMx systems while the Malay samples were genotyped on the Illumina HumanMap 6100Quad. We removed SNPs with call rate <95%, or departure from HWE (P<0.0001), or which were monomorphic.

Autopsy series. DNA was obtained using a modified DNA-extraction protocol [43] for genotyping using a Taqman kit from ABI and an ABI 7900HT Fast Real-Time PCR System.

Re-sequencing of the rs1408888 genomic region

The genomic region between rs1408888 and rs9572813 was PCR amplified in 2 DNA fragments for sequencing. The fragment close to rs1408888 was amplified by Dachi1-F (5'-TCTTGGTA-TAAAATGCAATGAAAGGAG-3') and 1R (5'-AAGGCGAAA-GGGAGGAAA-3'). The 1.7 Kb DNA fragment was sequenced by primers 1F (5'-AAGGGCCCATGACGAATG-3') and 3F (5'-TCACCTGAAATGAGTTACCA-3'). 30 traces were sequenced in one direction and 2R (5'-TTATTATCCGGGCCCCATTCC-3') in opposite direction. The primer 1F covers rs57143718 and the primer 3F covers CNE2003. The fragment close to rs9572813 was amplified by CNE2003-1F (5'-TACATACATCGCCCCAAAGGA-3') and Dachi1-R (5'-CAGAAATCTCCCACGGTGAC-3'). The fragment was sequenced by CNE2003-2F (5' -TGACCCAGCCTCCTGCTC -3') to cover rs17252745 and rs17252752. The DNA sequencing data were deposited to the NCBI Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?cmd = retrieve&val = CENTER_NAME%20%3D%20%22CUHK%22&dopt = info&size = 330&dispmax = 1&page = 1&recipes = Show), with the TI number 2333639769-2335840993 (300 traces).

Detection of DACH1 and CNES expression by reverse-transcription PCR

Expression of CNE2003 and DACHI in transcripts in various cell types were detected by RT-PCR. Expression of the CNE (220bp) was quantified by primers 5'-TATACATTGGCCCTAAGGA-3' and 5'-TTTGGATTTCAGCCTTGTCA-3'. Expression of Dachi1 was detected using 5'-CTGCAACGCGAAGTTGCT-3' and 5'-ATAAGCCTACGATCCCTG-3' as primers. Expression of β actin was used as a positive control using 5'-AGAGCTACGAGCTGCGTCG-3' and 5'-AGCACTGGTG-TGGCGTACAG-3' as primers. Expression of Dachi1 in PBMC was quantified by real-time PCR using the SYBR Green method with 5'-GTCGAAAACACCCCTCAGA-3' and 5' -CGAAGCTTCTCCTGGAGATG-3' as primers in an ABI 7900HT Real-Time PCR system. Expression level was normalized to the expression of β actin for comparison using the ΔΔCT method. The result was analyzed by Wilcoxon rank sum test and regression analysis with sex and age adjustment.

Statistical analysis

We used PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/), Statistical Analysis Software v.9.1 (SAS Institute, Cary, NC, USA) or Statistical Package for Social Sciences for Windows v.13 (SPSS, Chicago, IL, USA) for all statistical analyses, unless specified otherwise. All data are presented as mean±SD or median (interquartile range) unless specified. Categorical variables
were compared using χ² test, Fisher’s exact test and logistic regression, expressed as ORs and 95% CI as appropriate. In healthy controls, fasting plasma insulin, HOMA-IR and HOMA-β were logarithmically transformed due to skewed distributions. Genotype-phenotype associations were tested by multivariable linear regression adjusted for sex and age under the additive genetic model expressed in β coefficients with 95% CI. Between-group comparisons were performed by χ² test, Student’s t-test or Wilcoxon Rank Sum as appropriate. We used logistic regression to examine genetic association with CVD in a case-control cohort matched for age, sex and disease duration with further adjustment for logarithm of eGFR, expressed as OR (95% CI). A two-tailed P value <0.05 was considered significant.

Sample size estimation
Assuming an additive model with allele frequencies of 0.05–0.30, and an OR of 1.2–3.0 (for a prevalence of 0.1), we used the Genetic Power Calculator [45] to estimate the power for stage 1 (genome scan) and stage 2 (replication) at α levels of 1×10⁻⁴ and 0.05, respectively. For allele frequency >0.2, a sample size of 200 had >90% power to detect an OR of 3 and >75% power for an OR of 2.5. For the replication cohort with a sample size of 3000, we had >90% power to confirm an OR 1.2 for allele frequency >0.2. For the risk association with CVD, for SNP with allele frequency >0.2, a sample size of 2000 had over 90% power to confirm an OR of 1.5.

GWAS and meta-analysis
Distributions of all genotypes were analyzed for deviation from HWE by χ² test with one degree of freedom. In stage 1 experiment, we estimated possible familial relationship using estimates of identity-by-descent (IBD) derived from pair-wise analyses of 102,919 independent (r²<0.3) and quality SNPs. We did not detect population stratification using multidimensional scaling analysis and the inflation factor λ for GC, GC [46] was applied to correct for relatedness of the subjects and adjust for potential population stratification. The inflation factor λ was estimated by taking the median of the distribution of the χ² statistic from 425,513 quality SNPs in allelic test, and then divided by the median of the expected χ² distribution. The Quantile-Quantile plots were used to compare the observed and expected distributions for the 1df χ² statistics generated from allelic tests with or without correction for GC in the discovery stage. We calculated the corrected P values by dividing the observed χ² statistic by λ. For the top signals taken forward for replication, we used Haploview v4.1 to generate pairwise LD measures and the Manhattan plot as well as LocusZoom v1.1 to generate the regional plots for the interested gene loci. For analysis of data from stage 1, 2 and combined dataset, we used allelic χ² tests in 2×2 contingency tables to derive the OR after correction for multiple testing in 10,000 permutations. We used MIX v1.7 [47] to perform meta-analysis and calculated the combined estimates of ORs by weighting the natural log-transformed ORs (with respect to the same allele) of each study using the inverse of their variance under the fixed effect model. Cochran’s Q statistic (P<0.05) and I² were used to assess heterogeneity of ORs between studies.

Supporting Information
Table S1 Summary of bioinformatics analysis of rs1408888 of DACH1 (Ch13:70910099-71339331) based on NCBI Build 36.

Acknowledgments
We thank all study participants and health care professionals, especially Ms. Cherry Chiu and Ms. Rebecca Wong at the CUHK and Prince of Wales Hospital Diabetes and Endocrine Centre. We thank the CUHK Information Technology Services Centre and Centre for Clinical Trials for support of computing resources.

Dr. Juliana CN Chan is the guarantor of this work, had full access to all the data, and takes full responsibility for the integrity of data and the accuracy of data analysis.

Author Contributions
Conceived and designed the experiments: RCWM WYS MCYN JCNC. Performed the experiments: JG, J. Wong, HML, PSL, TCKL. Analyzed the data: RCWM, WYS, HML, CHTT, JSKH, YW, VL, EL, GZ, CHTT, HLZ, JG, HX, J. Wong, PSL, TCKL, JCNC. Wrote the paper: RCWM, VL, HML, CHTT, JSKH, JCNC. Bioinformatics and functional analyses: VI, EL, GZ, HML, CHTT, HLZ, JG, HX, J. Wong, PSL, TCKL. Researched data: CH, WPJ, KSP, HKL, HF, KN, EST, DPKN, NT, J. Woo, APSK, PCT, PCL. Critical revision of the manuscript: SKWT TFC AL. WYS GX CH KSP HF EST WJH HKL KN.

References
1. McCarroll MI (2010) Genomics, type 2 diabetes, and obesity. N Engl J Med 363: 2339–2350.
2. Cho YS, Chen CH, Hu C, Long J, Hee Ong RT, et al. (2011) Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet 44: 67–72.
3. Li H, Gao W, Lu L, Dong X, Han X, et al. (2013) A Genome-Wide Association Study Identified GRK5 and RASGRP1 as Type 2 Diabetes Loci in Chinese Hans. Diabetes 62: 291–298.
4. Shu XO, Long J, Qi L, Xiang YB, et al. (2010) Identification of new genetic risk variants for type 2 diabetes. PLoS Genet 6.
5. Ng MCY, Lee SC, Ko GTG, Li JKY, So WY, et al. (2001) Familial early onset type 2 diabetes in Chinese: the more significant roles of obesity and genetics than autoimmunity. Diabetes Care 24: 667–671.
6. Chan JC, Malik V, Ja W, Kadowaki T, Vajnik C, et al. (2009) Diabetes in Asia: epidemiology, risk factors, and pathophysiology. JAMA 301: 2129–2140.
7. Wang Y, Cai JH, Ma RCW, Song XY, Chan JCN, et al. (2013) Age of diagnosis, chronic hepatitis B virus infection and cardiovascular-renal endpoints in type 2 diabetes: a 10-year prospective cohort analysis by structural equation modeling. BMC Public Health: in press.
8. Li JK, Ng MC, So WY, Chiu CK, Otsuki R, et al. (2006) Phenotypic and genetic clustering of diabetes and metabolic syndrome in Chinese families with type 2 diabetes mellitus. Diabetes Metab Res Rev 22: 46–52.
9. Ng MC, So WY, Lam VK, Cockram CS, Bell GI, et al. (2004) Genome-wide scan for metabolic syndrome and related quantitative traits in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21–q25. Diabetes 53: 2676–2683.
10. Després JP, Lemieux I (2006) Abdominal obesity and metabolic syndrome. Nature 444: 881–887.
11. Ford ES (2005) Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence. Diabetes Care 28: 1769–1778.
12. Ma RC, Hui C, Tam CH, Zhang R, Kwan P, et al. (2013) Genome-wide Association Study in Chinese Identifies a Susceptibility Locus for Type 2 Diabetes at 7q32 near PAX4. Diabetologia 56: 1291–1305.
13. Kalousova A, Mavropoulos A, Adams BA, Nekrep N, Li Z, et al. (2010) Dachshund homologues play a conserved role in islet cell development. Dev Biol 340: 143–152.
14. Okamoto N, Nishimori Y, Nishimura T (2012) Conserved role for the Dachshund protein with Drosophila Pax6 homolog Eyeless in insulin expression. Proc Natl Acad Sci U S A 109: 2406–2411.
15. Nohregah MA, Ovcharenko I, Alafz V, Rubin EM (2003) Scanning human gene deserts for long-range enhancers. Science 302: 413.
16. Pennacchio LA, Ahn M, Moses AM, Prabhakar S, Nohregah MA, et al. (2006) In vivo enhancer analysis of human conserved non-coding sequences. Nature 444: 499–502.
17. Woolf A, Goodson M, Goodf OD, Smell P, McEwen GK, et al. (2005) Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol 3: e7.
18. Suen PM, Zou C, Zhang YA, Lau TK, Chan J, et al. (2008) PDZ-domain containing-2 (PDZD2) is a novel factor that affects the growth and differentiation of human fetal pancreatic progenitor cells. Int J Biochem Cell Biol 40: 789–803.

19. Wu K, Liu M, Li A, Denninger H, Rao M, et al. (2007) Cell fate determination factor DACH1 inhibits c-Jun-induced contact-independent growth. Mol Biol Cell 18: 755–767.

20. Zhou J, Wang C, Wang Z, Dampier W, Wu K, et al. (2010) Attenuation of Forkhead signaling by the retinal determination factor DACH1. Proc Natl Acad Sci U S A 107: 6861–6869.

21. Popov VM, Wu K, Zhou J, Powell MJ, Mardon G, et al. (2010) The Dachshund gene in development and hormone-responsive tumorigenesis. Trends Endocrinol Metab 21: 41–49.

22. Davis RJ, Shen W, Sandler YI, Amoui M, Purcell P, et al. (2001) Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality. Mol Cell Biol 21: 1484–1490.

23. Wu K, Yang Y, Wang C, Davoli MA, D'Anico M, et al. (2003) DACH1 inhibits transforming growth factor-beta signaling through binding Smad4. J Biol Chem 278: 51673–51684.

24. Ramos-Mondragon R, Galindo CA, Avila G (2008) Role of TGF-beta on smooth muscle: contribution to lipid binding and accelerated atherosclerosis in diabetes. J Diabetes 2: 233–242.

25. Zhou J, Wang C, Wang Z, Dampier W, Wu K, et al. (2010) Attenuation of Forkhead signaling by the retinal determination factor DACH1. Proc Natl Acad Sci U S A 107: 6861–6869.

26. Libby P, Theroux P (2005) Pathophysiology of coronary artery disease. Circulation 111: 3481–3488.

27. Sengupta U, Ukil S, Dimitrova N, Agrawal S (2009) Expression-based network biology identifies alteration in key regulatory pathways of type 2 diabetes and associated risk/complications. PLoS One 4: e8100.

28. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316: 1331–1336.

29. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, et al. (2007) Genome-wide association study identifies multiple loci that influence risk of type 2 diabetes. Science 316: 1336–1341.

30. Davis RJ, Shen W, Sandler YI, Amoui M, Purcell P, et al. (2001) Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality. Mol Cell Biol 21: 1484–1490.

31. Popov VM, Wu K, Zhou J, Powell MJ, Mardon G, et al. (2010) The Dachshund gene in development and hormone-responsive tumorigenesis. Trends Endocrinol Metab 21: 41–49.

32. Davis RJ, Shen W, Sandler YI, Amoui M, Purcell P, et al. (2001) Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality. Mol Cell Biol 21: 1484–1490.

33. Novembre J, Di Rienzo A (2009) Spatial patterns of variation due to natural selection in humans. Nat Rev Genet 10: 743–753.

34. Diamond J (2003) The double puzzle of diabetes. Nature 425: 599–602.

35. Fajans SS, Bell GI, Polsensky KS (2005) Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. N Engl J Med 354: 971–980.

36. Tan JT, Ng DP, Nurbaya S, Ye S, Lim XL, et al. (2010) Polymorphisms identified through genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore. J Clin Endocrinol Metab 95: 390–397.

37. Liu KH, Chan YL, Chan WB, Chan JC, Chu CW, et al. (2008) A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians. Diabetesologai 51: 602–608.

38. Hu C, Zhang R, Wang G, Ma X, Fang Q, et al. (2009) A genetic variant of G6PC2 is associated with type 2 diabetes and fasting plasma glucose level in the Chinese population. Diabetologia 52: 451–456.

39. Wang J, Liu M, Li A, Donninger H, Rao M, et al. (2007) Cell fate determination factor DACH1 inhibits c-Jun-induced contact-independent growth. Mol Biol Cell 18: 755–767.

40. Ng MC, Park KS, Oh B, Tam CH, Chuang LM, et al. (2008) A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians. Diabetesologai 51: 602–608.

41. Tan JT, Ng DP, Nurbaya S, Ye S, Lim XL, et al. (2010) Polymorphisms identified through genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore. J Clin Endocrinol Metab 95: 390–397.

42. Ng MC, Park KS, Oh B, Tam CH, Cho YM, et al. (2008) Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. Diabetes 57: 2226–2233.

43. Guan J, Zhao HL, Baum L, Sui Y, He L, et al. (2009) Apolipoprotein E polymorphism and expression in type 2 diabetic patients with nephropathy: clinicopathological correlation. Nephrol Dial Transplant 24: 1889–1895.

44. Zhao HL, Tong PC, Lai FM, Tomlinson B, Chan JC (2004) Association of the aldose reductase gene in development and hormone-responsive tumorigenesis. Trends Endocrinol Metab 95: 390–397.

45. Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore. J Clin Endocrinol Metab 95: 390–397.

46. Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997–1004.

47. Bax L, Yu LM, Ikeda N, Tsuruta H, Moons KG (2006) Development and validation of MIX: comprehensive free software for meta-analysis of causal research data. BMC Med Res Methodol 6: 50.