Published in final edited form as:
Nat Genet. ; 44(1): 67–72. doi:10.1038/ng.1019.

Meta-analysis of genome-wide association studies identifies 8 new loci for type 2 diabetes in East Asians

Yoon Shin Cho[1,¶,§], **Chien-Hsiun Chen**[2,3,§], **Cheng Hu**[4,§], **Jirong Long**[5,§], **Rick Twee Hee Ong**[6,§], **Xueling Sim**[7,§], **Fumihiko Takeuchi**[8,§], **Ying Wu**[9,§], **Min Jin Go**[1,§], **Toshimasa Yamauchi**[10,§], **Yi-Cheng Chang**[11,§], **Soo Heon Kwak**[12,§], **Ronald C.W. Ma**[13,§], **Ken Yamamoto**[14,§], **Linda S. Adair**[15], **Tin Aung**[16,17], **Qiuyin Cai**[5], **Li-Ching Chang**[2], **Yuan-Tsong Chen**[2], **Yutang Gao**[18], **Frank B. Hu**[19], **Hyung-Lae Kim**[1,20], **Sangsoo Kim**[21], **Young Jin Kim**[1], **Jeannette Jen-Mai Lee**[22], **Nanette R. Lee**[23], **Yun Li**[9,24], **Jian Jun Liu**[25], **Wei Lu**[26], **Jiro Nakamura**[27], **Eitaro Nakashima**[27,28], **Daniel Peng-Keat Ng**[22], **Wan Ting Tay**[16], **Fuu-Jen Tsai**[3], **Tien Yin Wong**[16,17,29], **Mitsuhiro Yokota**[30], **Wei Zheng**[5], **Rong Zhang**[4], **Congrong Wang**[4], **Wing Yee So**[13], **Keizo Ohnaka**[31], **Hiroshi Ikegami**[32], **Kazuo Hara**[10], **Young Min Cho**[12], **Nam H Cho**[33], **Tien-Jyun Chang**[11], **Yuqian Bao**[4], **Åsa K. Hedman**[34], **Andrew P. Morris**[34], **Mark I. McCarthy**[34,35], **DIAGRAM consortium**[36], **MuTHER consortium**[36], **Ryoichi Takayanagi**[31,*], **Kyong Soo Park**[12,37,*], **Weiping Jia**[4,*], **Lee-Ming Chuang**[11,38,*], **Juliana C.N. Chan**[13,*], **Shiro Maeda**[39,*], **Takashi Kadowaki**[10,*], **Jong-Young Lee**[1,*], **Jer-Yuarn Wu**[2,3,*], **Yik Ying Teo**[22,24,40,41,42,*], **E Shyong Tai**[43,44,45,¶,*], **Xiao Ou Shu**[5,*], **Karen L. Mohlke**[9,*], **Norihiro Kato**[8,*], **Bok-Ghee Han**[1,*], and **Mark Seielstad**[25,46,47,¶,*]

[1]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, The Republic of Korea

[2]Institute of Biomedical Sciences, Academia Sinica, 128 Academia Road Sec. 2, Nankang, Taipei, Taiwan

[3]School of Chinese Medicine, China Medical University, 91 Hsueh-Shih Road, Taichung, Taiwan

[4]Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, 600 Yishan Road, Shanghai, 200233, PR China

[5]Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

[6]NUS Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore 117456, Singapore

[¶]Correspondence should be addressed to: Y.S.C. (yooncho33@korea.kr), M.S. (Mark.Seielstad@ucsf.edu) or E.S.T. (eshyong@pacific.net.sg).
[36]A full list of members is provided in the Supplementary Note.
[§]These authors contributed equally as co-first authors.
[*]These authors jointly supervised this work.

AUTHOR CONTRIBUTIONS

The study was supervised by E.S.T., B.G.H., N.K., Y.S.C., Y.Y.T., W.Z., Q.C., X.O.S., Y.-T.C., J.-Y.W., L.S.A., K.L.M., T.K., C.H., W.J., L.-M.C., Y.M.C., K.S.P., J.-Y.L. and J.C. The experiments were conceived and designed by Y.S.C., E.S.T., N.K., D.P.-K.N., J.J.-M.L., M.S., T.Y.W., Y.Y.T., W.Z., F.H., X.O.S., C.-H.C., F.-J.T., Y.-T.C., J.-Y.W., L.S.A., K.L.M., S.M., C.H., L.-M.C., K.S.P., M.J.G. M.I.M. and R.M. The experiments were performed by J.L., M.S., J.L., J.-Y.W., S.M., R.Z., K.Y., Y.C., T.-J.C., L.-M.C. and S.H.K. Statistical analysis was performed by M.J.G., X.S., Y.J.K., R.T.H.O., W.T.T., Y.Y.T., F.T., J.L., C.-H.C., L.-C.C., Y.W., Y.L., K.H., C.H., Y.C., S.H.K. A.P.M. and R.M. The data were analyzed by M.J.G., X.S., Y.J.K., R.T.H.O., W.T.T., Y.Y.T., J.L., C.-H.C., L.-C.C., Y.W, N.R.L., Y.L., L.S.A., K.L.M., T.Y., C.H., Y.C., S.H.K., Y.S.C., S.K. A.K.H. and R.M. Reagents, materials and analysis tools were contributed by E.S.T., B.G.H., N.K., D.P.-K.N., J.J.-M.L., J.L., M.S., T.A., T.Y.W., E.N., M.Y., J.N., J.L., W.Z., Q.C., Y.G., W.L., F.B.H., X.O.S., F.-J.T., Y.-T.C., J.-Y.W., N.R.L., Y.L., K.O., H.I., R.T., C.W., Y.B., T.-J.C, L.-M.C., K.S.P., H.-L.K., N.H.C., J.-Y.L., W.Y.S. and J.C. The manuscript was written by Y.S.C., M.S. and E.S.T. All authors reviewed the manuscript.

[7]Centre for Molecular Epidemiology, National University of Singapore, Singapore 117597, Singapore

[8]Research Institute, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, 162-8655, JAPAN

[9]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA

[10]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

[11]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[12]Department of Internal Medicine, Seoul National University College of Medicine, 101 Daehak-Ro, Jongno-Gu, Seoul, 110-744, Korea

[13]Dept of Medicine and Therapeutics, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong

[14]Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582 Japan

[15]Department of Nutrition, University of North Carolina, Chapel Hill, NC, USA

[16]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore 168751, Singapore

[17]Department of Ophthalmology, National University of Singapore, Singapore 119074, Singapore

[18]Department of Epidemiology, Shanghai Cancer Institute, Shanghai 200032, China

[19]Department of Nutrition and Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America

[20]Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, The Republic of Korea

[21]School of Systems Biomedical Science, Soongsil University, Sangdo-5-dong 1-1, Dongjak-gu, Seoul 156-743, The Republic of Korea

[22]Department of Epidemiology and Public Health, National University of Singapore, Singapore 117597, Singapore

[23]Office of Population Studies Foundation Inc., University of San Carlos, Cebu City, Philippines

[24]Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599

[25]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138672, Singapore

[26]Shanghai Institute of Preventive Medicine, Shanghai 200336, China

[27]Division of Endocrinology and Diabetes, Department of Internal Medicine, Nagoya University Graduate School of Medicine, Nagoya, 466-8560 JAPAN

[28]Department of Diabetes and Endocrinology, Chubu Rosai Hospital, Nagoya, 455-8530 Japan

[29]Centre for Eye Research Australia, University of Melbourne, East Melbourne VIC, 3002 Australia

[30]Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya, 464-8651 Japan

[31]Department of Geriatric Medicine, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582 Japan

[32]Dept Endocrinology, Metabolism and Diabetes, Kinki University School of Medicine 377-2 Ohno-higashi, Osaka-sayama, Osaka, 589-8511 Japan

[33]Department of Preventive Medicine, Ajou University School of Medicine, Suwon, The Republic of Korea

[34]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

[35]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, UK

[37]WCU Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology and College of Medicine, Seoul National University, 101 Daehak-Ro, Jongno-Gu, Seoul, 110-744, Korea

[38]Graduate Institute of Clinical Medicine, National Taiwan University School of Medicine, Taipei, Taiwan

[39]Laboratory for Endocrinology and Metabolism, RIKEN Center for Genomic Medicine, Yokohama 230-0045, Japan

[40]Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546, Singapore

[41]Centre for Molecular Epidemiology, National University of Singapore, Singapore 117597, Singapore

[42]NUS Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore 117456, Singapore

[43]Department of Medicine, National University of Singapore, Singapore 119228 Singapore

[44]Department of Epidemiology and Public Health, National University of Singapore, Singapore 117597, Singapore

[45]Duke-National University of Singapore Graduate Medical School, Singapore 169857, Singapore

[46]Institute for Human Genetics, University of California, 513 Parnassus Avenue, San Francisco, CA 94143-0794, USA

[47]Blood Systems Research Institute, 270 Masonic Avenue, San Francisco, California, 94118, USA

Abstract

We conducted a three-stage genetic study to identify susceptibility loci for type 2 diabetes (T2D) in East Asian populations. The first stage meta-analysis of eight T2D genome-wide association studies (6,952 cases and 11,865 controls) was followed by a second stage *in silico* replication analysis (5,843 cases and 4,574 controls) and a stage 3 *de novo* replication analysis (12,284 cases and 13,172 controls). The combined analysis identified eight new T2D loci reaching genome-wide significance, which were mapped in or near *GLIS3*, *PEPD*, *FITM2-R3HDML-HNF4A*, *KCNK16*, *MAEA*, *GCC1-PAX4*, *PSMD6* and *ZFAND3. GLIS3*, involved in pancreatic beta cell development and insulin gene expression[1,2], is known for its association with fasting glucose levels[3,4]. The evidence of T2D association for *PEPD*[5] and *HNF4A*[6,7] has been detected in previous studies. *KCNK16* may regulate glucose-dependent insulin secretion in the pancreas. These findings derived from East Asians provide new perspectives on the etiology of T2D.

Type 2 diabetes (T2D) is a major public health problem whose global prevalence is rapidly rising[8]. The development of T2D is influenced by diverse factors, and decades of epidemiological studies have linked obesity, hypertension, and dyslipidemia with the risk of

T2D[9]. It is also known that T2D shows considerable heritability. Within only the last three years, genetic studies have yielded a rapidly lengthening list of loci harboring disease predisposing variations[10]. To date, genetic variants at forty-five loci have been identified for T2D[10,11]. Despite these advances toward a better understanding of the genetic basis of T2D, its heritability is yet to be fully explained[12]. In addition, most T2D loci have been detected initially in population samples of European origin, apart from *KCNQ1*, *UBE2E2*, and *C2CD4A-C2CD4B*, which were first identified in East Asian studies[13–15]. Additional efforts involving East Asian populations have identified variants at the *SPRY2*, *PTPRD* and *SRR* loci[5,16,17]. However, these have not been extensively replicated in multiple populations. A large meta-analysis in East Asians is expected to identify novel genetic associations and provide insights into T2D pathogenesis. In addition to differences in the allele frequencies between East Asians and Europeans, which may affect the power to detect associations, T2D epidemiology also differs considerably among different European populations and East Asian populations. In East Asians, the rates of diabetes are often higher at a lower average BMI's[18], suggesting that some different pathways may be involved in pathogenesis of T2D.

To discover new T2D loci, we conducted a three-stage association study in individuals of East Asian descent (Supplementary Fig. 1). The stage 1 meta-analysis combining 8 T2D GWA studies participating in the Asian Genetic Epidemiology Network (AGEN) consortium (6,952 cases and 11,865 controls) was performed using association data for 2,626,356 imputed and genotyped autosomal SNPs by the inverse-variance method for fixed effects (Supplementary Table 1). All imputed and genotyped SNPs (minor allele frequency > 0.01) passed quality control filters in each of the eight stage 1 data sets prior to conducting meta-analysis (Supplementary Table 2). The genomic control inflation factor (λ) for the meta-analysis was 1.046 (less than 1.062 for the individual studies), indicating that results in stage 1 were not likely resulting from population stratification (Supplementary Fig. 2). Individuals from each component study participated in stage 1 mostly clustered together with CHB/JPT HapMap samples in the principal component analysis plot (Supplementary Fig. 3), further demonstrating the similarity in the ethnicity between stage 1 samples. Most signals showing strong evidence for T2D associations appeared in previously known T2D genes (Fig. 1). Stage 1 *P*-values, OR's, and average risk allele frequencies for 45 previously reported T2D associated SNPs are shown in Supplementary Table 3.

After removing known T2D variants, 297 SNPs from independent loci were selected from the stage 1 meta-analysis based on our arbitrary inclusion criteria for follow-up *in silico* replication: meta-analysis *P*-value $< 5\times10^{-4}$ (based on the divergence between the observed and expected *P*-values on the Q-Q plot (Supplementary Fig. 2)), heterogeneity *P*-value>0.01 and the number of studies included in meta-analysis ≥ 7 (Supplementary Table 4). A total of 3,756 SNPs including the 297 selected SNPs and their proxies (r^2 >0.8 based on phase2 CHB/JPT HapMap data) were taken forward to stage 2, an *in silico* replication, in three independent GWA studies (5,843 cases and 4,574 controls). After meta-analysis combining stage 1 and stage 2 data for 3,756 SNPs, we selected the 19 SNPs showing the most compelling evidence for association (stage 1 and 2 combined *P*-value$<10^{-5}$) (Supplementary Table 5) for stage 3 *de novo* genotyping in up to 12,284 cases and 13,172 controls recruited from five independent studies (Supplementary Tables 1 & 2). This resulted in 8 novel T2D loci reaching genome-wide significance from the combined meta-analysis across all three stages (Table 1 & Fig. 2).

Three of these eight loci have been associated previously with metabolic traits or related diseases, or suggestively with T2D itself. One such locus was detected within an intron of *GLIS3*, a gene that is highly expressed in islet beta cells. The coding product of this gene, a Krüppel-like zinc finger transcription factor, has been proposed as a critical player in the regulation of pancreatic beta cell development and insulin gene expression[1,2]. SNPs in high

LD with this locus have already been implicated in association with type 1 diabetes (T1D)[19] and fasting plasma glucose[3]. The second locus, on 19q13, is located in an intron of *PEPD*. Several SNPs (lead SNP: rs10425678) in this gene were identified in association with T2D previously in Japanese [5]. However, the SNP in our study (rs3786897) is not in LD with those identified in Japanese (r^2 = 0.008 and D′ = 0.143 between rs3786897 and rs10425678), and our GWA data do not support an association for T2D with rs10425678 (*P* = 0.528). The third such signal is near *FITM2-R3HDML-HNF4A*. *FITM2* may be involved in lipid droplet accumulation[20], while the function of *R3HDML* is not known. Mutations in *HNF4A* cause maturity onset diabetes of the young type 1 (MODY1)[21]. Common variants, rs1884613 and rs2144908, in the P2 promoter region of this gene have been associated with T2D in a population-specific manner[6,22]. The SNPs in our study (rs6017317) are not in strong LD with *HNF4A* P2 promoter SNPs (r^2 = 0.23~0.25, D′ = 0.50~0.54 based on phase2 CHB/JPT HapMap data), indicating that rs6017317 is a new T2D signal in the 20q13.12 region where *HNF4A* resides.

The remaining five loci reaching genome-wide significance in our study have not previously been reported in the context of any metabolic traits, including the loci mapped in or near *KCNK16*, *MAEA*, *GCC1-PAX4*, *PSMD6* and *ZFAND3*. *KCNK16*, expressed predominantly in the pancreas, encodes a potassium channel protein containing two pore-forming P domains[23]. In pancreatic β cells, potassium channels that are inhibited by ATP regulate glucose-dependent insulin secretion. Among the variants in strong LD with this signal (rs1535500) is rs11756091 (r^2 = 0.977, D′ = 1.0 based on phase2 CHB/JPT HapMap data), which encodes a proline-to-histidine substitution in two isoforms of KCNK16. This or other variants influencing this gene may result in defective regulation of potassium channel activity contributing to the etiology of T2D[24]. *MAEA* encodes a protein that plays a role in erythroblast enucleation and in the development of mature macrophages[25]. The gene-set analysis of the stage 1 *P*-values using GSA-SNP[26] indicated that *MAEA* belongs to a group of genes that showed significant association with T2D and includes *IDE* located at a known T2D susceptibility locus[27] (stage 1 $P = 1.41\times10^{-7}$ for rs6583826 at the *IDE* locus in this study). The GRIP domain containing protein encoded by *GCC1* may play a role in the organization of the trans-Golgi network involved in membrane transport [28]. *PAX4* that is only 30kb further away from *GCC1* is an outstanding candidate for T2D given its involvement in pancreatic islet development. This gene was recently implicated in a Japanese case of MODY[29]. The expression product of *PSMD6* that acts as a regulatory subunit of the 26S proteasome is likely involved in the ATP-dependent degradation of ubiquitinated proteins [30]. Although the function of *ZFAND3* has not been fully elucidated, it is noteworthy that family member *ZFAND6* is present along with *FAH* at a T2D locus detected previously[31]. We examined whether 8 novel loci may be associated with T2D through an effect on obesity, as seen with *FTO*[32]. All T2D association signals remained after the adjustment with BMI (Supplementary Table 6), indicating that the association with T2D for 8 loci are not mediated through an effect on obesity.

In addition to 8 loci reaching genome-wide significance, we identified 2 loci showing moderate evidence (combined *P*-value<10^{-6}) of association with T2D including *WWOX* and *CMIP* (Table 1). We looked up the association results for these 10 loci in GWA data from up to 47,117 European samples generated by the DIAGRAM consortium (‘DIAGRAM +’ for the current version of data set)[31]. Results from lookup for these loci indicated that three loci including *FITM2-R3HDML-HNF4A* (rs6017317: $P = 1.47\times10^{-2}$, OR = 1.07), *CMIP* (rs16955379: $P = 3.33\times10^{-2}$, OR = 1.20] and *MAEA* (proxy SNP rs11247991 for rs6815464 (r^2=0.96): $P = 6.56\times10^{-3}$, OR = 1.19) were modestly associated, while a locus in *GLIS3* (rs7041847: $P = 6.43\times10^{-2}$, OR = 1.04) was nominally associated with T2D. The direction of effect was consistent in four (*PSMD6*, *PEPD*, *WWOX* and *KCNK16*) of six loci that were not replicated in DIAGRAM+ (Supplementary Table 5).

The functional connections among 10 new T2D genes and 28 known T2D genes that were replicated in this study (Supplementary Table 3) were analyzed by GRAIL[33]. Connection results highlighted notable biological functions of sets of genes within T2D-assocaited regions (Supplementary Fig. 4 & Supplementary Tables 7 & 8). For example, *KCNK16* shows strong connections with previously known T2D genes encoding potassium channels (*KCNJ11* and *KCNQ1*), implying its physiological role in the regulation of potassium transport in the pancreatic cells.

The association between each new T2D SNP and the expression levels of genes within 1Mb of the SNP was examined by eQTL analysis in the MuTHER consortium. One SNP (rs3786897) in an intron of *PEPD* was highly associated with mRNA expression levels of *PEPD* in the adipose tissue of 776 individuals of European ancestry ($P_{eQTL} = 2.14\times10^{-8}$) (Supplementary Table 9). However, this SNP did not show an association with T2D in populations of European ancestry, and the significance of this finding is unclear.

We considered the possibility that autoimmune diabetes may be driving some of the signals observed. Firstly, cases in all studies were predominantly those with adult onset diabetes (age of disease onset ≥ 30 years), and none of the clinically diagnosed patients had T1D, defined as acute ketosis and continuous requirement of insulin within 1 year of diagnosis. Secondly, we looked up the associations for all known T1D associated variants in our study. Only a given number of loci survived from the tests applied (Supplementary Table 10). This is in distinct contrast to the situation for known T2D associated variants where many replicated in our study (Supplementary Table 3), further suggesting that our findings are most relevant to T2D. Finally, since variants close to the *GLIS3* locus have been shown to be associated with T1D[19], we examined the association between rs7041847 and diabetes in four studies (n=8,383) in which cases with positive GAD antibodies had been excluded (data not shown). In each study, the association between this SNP and diabetes was the same as when all samples were included (meta-analysis *P*-value = 3.4×10^{-4}, OR = 1.12). This finding, along with the fact that SNPs at the *GLIS3* locus also show associations with fasting plasma glucose in nondiabetic adults[3] and healthy children and adolescents[4], is consistent with the hypothesis that SNPs at this locus may impact fasting glucose homeostasis rather than the immune system. Taken together, it is unlikely that a significant proportion of the positive associations observed in our study were driven by autoimmune diabetes.

This study is the largest GWA meta-analysis conducted for T2D in East Asians. Findings from this study highlight not only previously unknown biological pathways but also population specific loci for T2D. The association of rs9470794 in *ZFAND3* with T2D looks highly specific to East Asians (Supplementary Table 5), whereas rs11634397 near *ZFAND6* looks specific to Europeans (Supplementary Table 3). A significant difference in risk allele frequencies of both loci is observed between the two continental populations (rs9470794 RAF = 0.32 CHB/JPT vs 0.12 CEU; rs11634397 RAF = 0.07 CHB/JPT vs 0.64 CEU). Although these loci are related to T2D differently in the two populations, these results lead one to speculate that the broader A20 domain-containing zinc finger protein family plays a role in the etiology of T2D. Additional population specific T2D loci are further suggested, as seen in *WWOX* (rs17797882) (Supplementary Table 5) for East Asians and *ZBED3* (rs4457053) (Supplementary Table 3) for Europeans. Despite the lack of clear physiological evidence, these findings may provide clues to understand population characteristic T2D phenotypes as exemplified by the high rates of diabetes at lower average BMI's in East Asian and other populations.

ONLINE METHODS

Study subjects

Stage 1 subjects were drawn from 8 T2D GWA studies participating in the AGEN consortium that was organized for genetic studies on diverse complex traits in 2010. These eight studies included 6,952 T2D cases and 11,865 controls from the Korea Association Resource Study (KARE), the Singapore Diabetes Cohort Study (SDCS), the Singapore Prospective Study Program (SP2), the Singapore Malay Eye Study (SiMES), the Japan Cardiometabolic Genome Epidemiology Network(CAGE), the Shanghai Diabetes Genetic Study (SDGS), the Taiwan T2D Study (TDS), and the Cebu Longitudinal Health and Nutritional Survey (CLHNS). Stage 2 subjects included 5,843 cases and 4,574 controls from three independent GWA studies including the BioBank Japan Study (BBJ), the Health2 T2D Study (H2T2DS) and the Shanghai Jiao Tong University Diabetes Study (SJTUDS) for *in silico* replication analysis. Stage 3 included up to, 12,284 cases and 13,172 controls from five different studies comprised of the Japan Cardiometabolic Genome Epidemiology Network (CAGE), the Shanghai Diabetes Study I/II (SDS I/II), the Chinese University of Hong Kong Diabetes Study (CUHKDS), the National Taiwan University Hospital Diabetes Study (NTUHDS) and the Seoul National University Hospital Diabetes Study (SNUHDS) for *de novo* replication analysis. The study design and T2D diagnosis criteria of each study in stages 1, 2, and 3 are described in Supplementary Table 1 and Supplementary Note. Each study obtained approval from the appropriate institutional review board, and written informed consent from all participants. The three-stage design of the overall study is depicted in Supplementary Fig. 1.

Genotyping and imputation

Subjects for stage 1 and 2 analyses were genotyped with high-density SNP typing platforms covering the entire human genome. In most studies, only unrelated samples with missing genotype call-rates below 5% were included for subsequent GWA analyses. For the genome-wide association meta-analysis, each study participating in stages 1 and 2 performed SNP imputation. IMPUTE, MACH or BEAGLE were used, together with haplotype reference panels from the JPT and CHB founders (JPT+CHB+CEU and/or YRI in some studies) on the basis of HapMap build 36 (release 21, 22, 23a or 24). Only imputed SNPs with high genotype information content (proper info > 0.5 for IMPUTE and Rsq> 0.3 for MACH and BEAGLE) were used for association analysis. Genotyping for the stage 3 analysis was carried out by TaqMan, Sequenom MassARRAY, or the Beckman SNP Stream method. All SNPs included in stage 3 satisfied a genotype success rate over 98% (Supplementary Table 2).

Statistical analyses, analysis tools, and SNP prioritization for stage 2 and 3

Associations between SNPs and T2D were tested by logistic regression with an additive model (1-d.f.) after adjustment for sex. Other adjustments were permitted according to the situation of individual studies. Meta-analysis was performed by an inverse-variance method assuming fixed effects with Cochran's Q test to assess between-study heterogeneity. METAL software (http://www.sph.umich.edu/csg/abecasis/Metal) was used for all meta-analyses. A plot of the negative log of association results from the stage 1 meta-analysis, by chromosome, was generated by the WGAViewer software (http://people.genome.duke.edu/~dg48/WGAViewer/std.php). The quantile-quantile plot was constructed by plotting the distribution of observed *P*-values of given SNPs against the theoretical distribution of expected *P*-values for T2D[34]. The genomic control inflation factor, λ, was calculated by dividing median χ^2 statistics by 0.456[35] for individual GWA studies as well as stage 1 GWA meta-analysis. We did not correct for genomic control in the stage 1 analyses as inflation was modest, suggesting that population structure is unlikely to cause significant

inflation of stage 1 results (Supplementary Table 2). Selection criteria for lead SNPs to take forward to stage 2 *in silico* replication analysis were (1) stage 1 meta-analysis *P*-value < 5×10^{-4} (based on the divergence between the observed and expected *P*-values on the Q-Q plot (Supplementary Fig. 2)), (2) heterogeneity *P*-value > 0.01 and (3) number of studies included in stage 1 meta-analysis ≥ 7 (Supplementary Table 4). Removing known variants associated with T2D proxies for each lead SNP ($r^2 > 0.8$) were selected using the SNAP software (http://www.broadinstitute.org/mpg/snap/)-HapMap. Replication genotyping for stage 3 was performed for novel SNPs with stage 2 combined *P*-value < 10^{-5}. Regional association results from genome-wide meta-analysis were plotted by the LocusZoom software (http://csg.sph.umich.edu/locuszoom/) for SNPs reaching genome-wide significance from combined meta-analysis of stage 1, 2 and 3.

Principal components analysis

A list of 76,534 common SNPs across Illumina 550/610/1M and Affymetrix 5.0/6.0 were first selected. These set of SNPs on the Asian HapMap II samples CHB+JPT were then trained to generate a list of 44,524 SNPs with pairwise LD < 0.3 in a window side of 50 SNPs. Individuals from each component study and HapMap II were plotted based on the first two eigenvectors produced by PCA.

Expression quantitative trait locus (eQTL) analysis

Gene expression information in 776 adipose tissues, 667 skin tissues and 777 LCLs was obtained from the MuTHER consortium[36]. The eQTL data for 8 of 10 T2D loci identified in this study were filtered at MAF > 5% and INFO > 0.8 except rs16955379 that has a MAF of 1.5% in the MuTHER data set. Two loci, rs6815464 (chr. 4) and rs17797882 (chr. 16), are not included in the MuTHER data set. Association between each significant SNP for T2D and normalized mRNA expression values of genes within 1Mb of the lead SNP were performed with the GenABEL/ProbABEL package using the polygenic linear model incorporating a kinship matrix in GenABEL followed by the ProbABEL mmscore test with imputed genotypes. A multiple-testing correction was applied to the *cis*-association results. *P*-value thresholds of 5.06×10^{-5} in adipose, 3.81×10^{-5} in skin and 7.80×10^{-5} in LCL correspond to an estimated genome-wide FDR of 1%.

Gene relationships among implicated loci (GRAIL) analysis

A GRAIL analysis was performed as described previously[31,33]. A total of 38 genes within T2D-associated regions were selected for the analysis. Among them, 28 genes were from the previously implicated set (Supplementary Table 3), while the other 10 were from ones newly implicated in this study (Table 1). PubMed abstracts published after December, 2006 were omitted from the analysis to reduce confounding by results from T2D GWA studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all the participants and the staff of the BioBank Japan project. The project was supported by a grant from the Leading Project of Ministry of Education, Culture, Sports, Science and Technology Japan.

CAGE: The CAGE Network Studies were supported by grants for the Program for Promotion of Fundamental Studies in Health Sciences, National Institute of Biomedical Innovation Organization (NIBIO); the Core Research for Evolutional Science and Technology (CREST) from the Japan Science Technology Agency; and the Grant of National Center for Global Health and Medicine (NCGM).

CLHNS: We thank the Office of Population Studies Foundation research and data collection teams for the Cebu Longitudinal Health and Nutrition Survey. This work was supported by National Institutes of Health grants DK078150, TW05596, HL085144, TW008288 and pilot funds from RR20649, ES10126, and DK56350.

CUHK: We acknowledge support from the Hong Kong Government Research Grants Council Central Allocation Scheme (CUHK 1/04C), Research Grants Council Earmarked Research Grant (CUHK4724/07M) and the Innovation and Technology Fund of the Government of the Hong Kong SAR (ITS/487/09FP). We acknowledge the Chinese University of Hong Kong Information Technology Services Center for support of computing resources. We would also like to thank the dedicated medical and nursing staff at the Prince of Wales Hospital Diabetes and Endocrine Centre.

KNIH: This work was supported by grants from Korea Centers for Disease Control and Prevention (4845-301, 4851-302, 4851-307) and an intramural grant from the Korea National Institute of Health (2010-N73002-00), the Republic of Korea.

NTUH: This work was supported in part by the grant (NSC99-3112-B-002-019) from the National Science Council of Taiwan. We would also like to acknowledge the National Genotyping Center of National Research Program for Genomic Medicine (NSC98-3112-B-001-037), Taiwan.

SDGS: This work was supported in part by US National Institutes of Health grants R01CA124558, R01CA64277, R01CA70867, R01CA90899, R01CA100374, R01CA118229, R01CA92585, UL1 RR024975, DK58845 and HG004399, DOD Idea Award BC050791, Vanderbilt Ingram professorship funds and Allen Foundation Fund. We would like to thank the dedicated investigators and staff members from research teams at Vanderbilt University, Shanghai Cancer Institute, and Shanghai Institute of Preventive Medicine, and most of all, the study participants for their contributions in the studies.

SDS: This work was supported by grants from National 973 Program (2011CB504001), Project of National Natural Science Foundation of China (30800617) and Shanghai Rising-Star Program (09QA1404400), China.

SJTUDS: This work was supported by grants from National 863 Program (2006AA02A409) and major program of Shanghai Municipality for Basic Research (08dj1400601), China

SNUH: This work was supported by grants from the Korea Health 21 R & D Project, Ministry of Health & Welfare (00-PJ3-PG6-GN07-001) and WCU project of the MEST and NRF(R31-2008-000-10103-0), Korea

References

1. Kang HS, et al. Transcription factor Glis3, a novel critical player in the regulation of pancreatic beta-cell development and insulin gene expression. Mol Cell Biol. 2009; 29:6366–79. [PubMed: 19805515]
2. Yang Y, Chang BH, Samson SL, Li MV, Chan L. The Kruppel-like zinc finger protein Glis3 directly and indirectly activates insulin gene transcription. Nucleic Acids Res. 2009; 37:2529–38. [PubMed: 19264802]
3. Dupuis J, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010; 42:105–16. [PubMed: 20081858]
4. Barker A, et al. Association of genetic Loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. Diabetes. 2011; 60:1805–12. [PubMed: 21515849]
5. 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. 2009; 58:1690–9. [PubMed: 19401414]
6. Barroso I, et al. Population-specific risk of type 2 diabetes conferred by HNF4A P2 promoter variants: a lesson for replication studies. Diabetes. 2008; 57:3161–5. [PubMed: 18728231]
7. Silander K, et al. Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. Diabetes. 2004; 53:1141–9. [PubMed: 15047633]
8. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. Nature. 2001; 414:782–7. [PubMed: 11742409]
9. Tkac I. Metabolic syndrome in relationship to type 2 diabetes and atherosclerosis. Diabetes Res Clin Pract. 2005; 68 (Suppl1):S2–9. [PubMed: 15955370]
10. Prokopenko I, McCarthy MI, Lindgren CM. Type 2 diabetes: new genes, new understanding. Trends Genet. 2008; 24:613–21. [PubMed: 18952314]

11. Rung J, et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet. 2009; 41:1110–5. [PubMed: 19734900]
12. Manolio TA, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461:747–53. [PubMed: 19812666]
13. Yasuda K, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet. 2008; 40:1092–7. [PubMed: 18711367]
14. Unoki H, et al. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet. 2008; 40:1098–102. [PubMed: 18711366]
15. Yamauchi T, et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at UBE2E2 and C2CD4A-C2CD4B. Nat Genet. 42:864–8. [PubMed: 20818381]
16. Tsai FJ, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet. 2010; 6:e1000847. [PubMed: 20174558]
17. Shu XO, et al. Identification of new genetic risk variants for type 2 diabetes. PLoS Genet. 2010; 6
18. Stommel M, Schoenborn CA. Variations in BMI and prevalence of health risks in diverse racial and ethnic populations. Obesity (Silver Spring). 2010; 18:1821–6. [PubMed: 20075855]
19. Barrett JC, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009; 41:703–7. [PubMed: 19430480]
20. Kadereit B, et al. Evolutionarily conserved gene family important for fat storage. Proc Natl Acad Sci U S A. 2008; 105:94–9. [PubMed: 18160536]
21. Nakajima H, et al. Hepatocyte nuclear factor-4 alpha gene mutations in Japanese non-insulin dependent diabetes mellitus (NIDDM) patients. Res Commun Mol Pathol Pharmacol. 1996; 94:327–30. [PubMed: 9029679]
22. Johansson S, et al. Studies in 3,523 Norwegians and meta-analysis in 11,571 subjects indicate that variants in the hepatocyte nuclear factor 4 alpha (HNF4A) P2 region are associated with type 2 diabetes in Scandinavians. Diabetes. 2007; 56:3112–7. [PubMed: 17827402]
23. Girard C, et al. Genomic and functional characteristics of novel human pancreatic 2P domain K(+) channels. Biochem Biophys Res Commun. 2001; 282:249–56. [PubMed: 11263999]
24. Ashcroft FM. ATP-sensitive potassium channelopathies: focus on insulin secretion. J Clin Invest. 2005; 115:2047–58. [PubMed: 16075046]
25. Soni S, et al. Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. J Biol Chem. 2006; 281:20181–9. [PubMed: 16707498]
26. Nam D, Kim J, Kim SY, Kim S. GSA-SNP: a general approach for gene set analysis of polymorphisms. Nucleic Acids Res. 2010; 38:W749–54. [PubMed: 20501604]
27. Scott LJ, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007; 316:1341–5. [PubMed: 17463248]
28. Luke MR, Houghton F, Perugini MA, Gleeson PA. The trans-Golgi network GRIP-domain proteins form alpha-helical homodimers. Biochem J. 2005; 388:835–41. [PubMed: 15654769]
29. Jo W, Endo M, Ishizu K, Nakamura A, Tajima T. A novel PAX4 mutation in a Japanese patient with maturity-onset diabetes of the young. Tohoku J Exp Med. 2011; 223:113–8. [PubMed: 21263211]
30. Wang X, et al. Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. Biochemistry. 2007; 46:3553–65. [PubMed: 17323924]
31. Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010; 42:579–89. [PubMed: 20581827]
32. Frayling TM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007; 316:889–94. [PubMed: 17434869]
33. Raychaudhuri S, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet. 2009; 5:e1000534. [PubMed: 19557189]
34. Hyndman RJ, Fan Y. Sample quantiles in statistical packages. Amer Stat. 1996; 50:361–65.
35. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. Theor Popul Biol. 2001; 60:155–66. [PubMed: 11855950]

36. Nica AC, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet. 2011; 7:e1002003. [PubMed: 21304890]

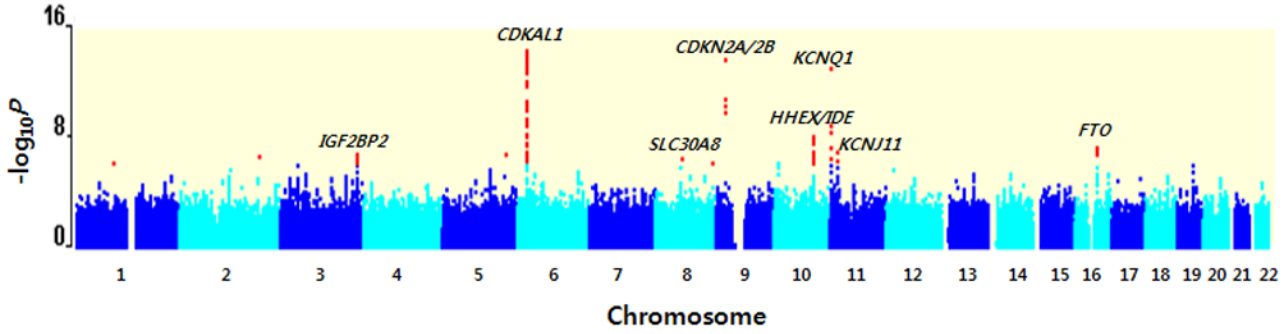
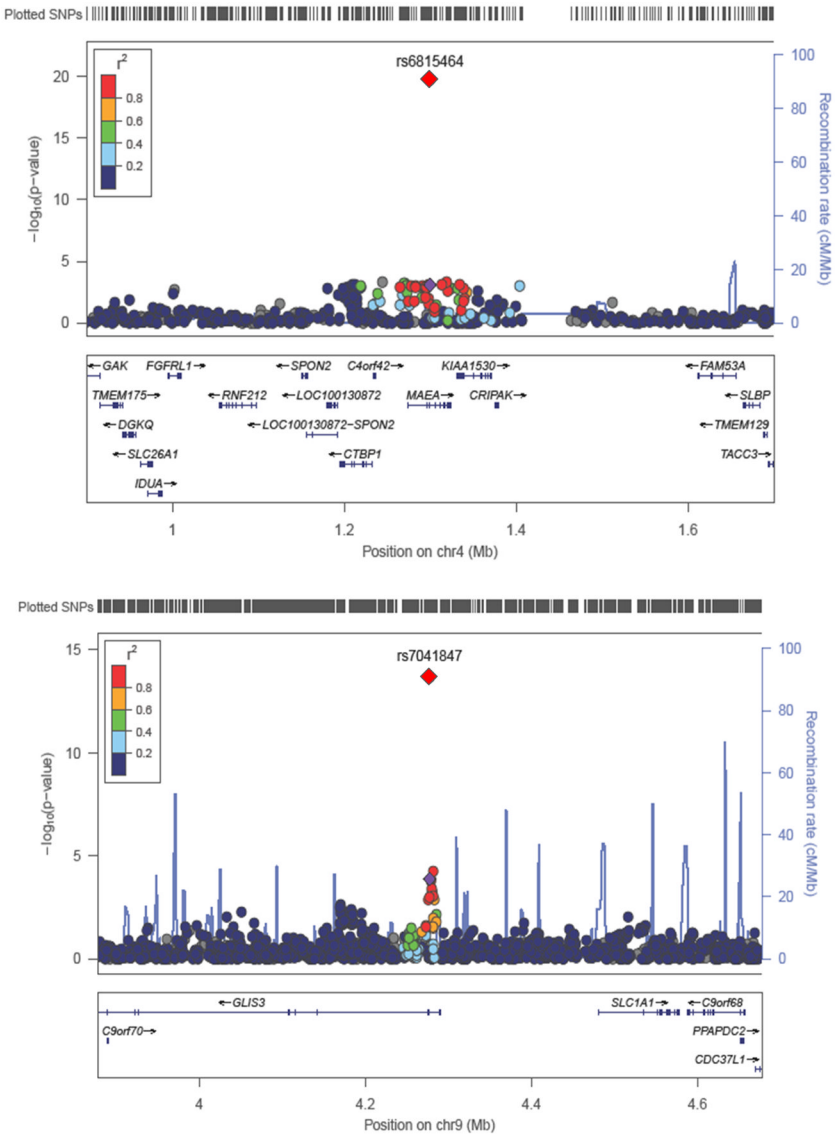
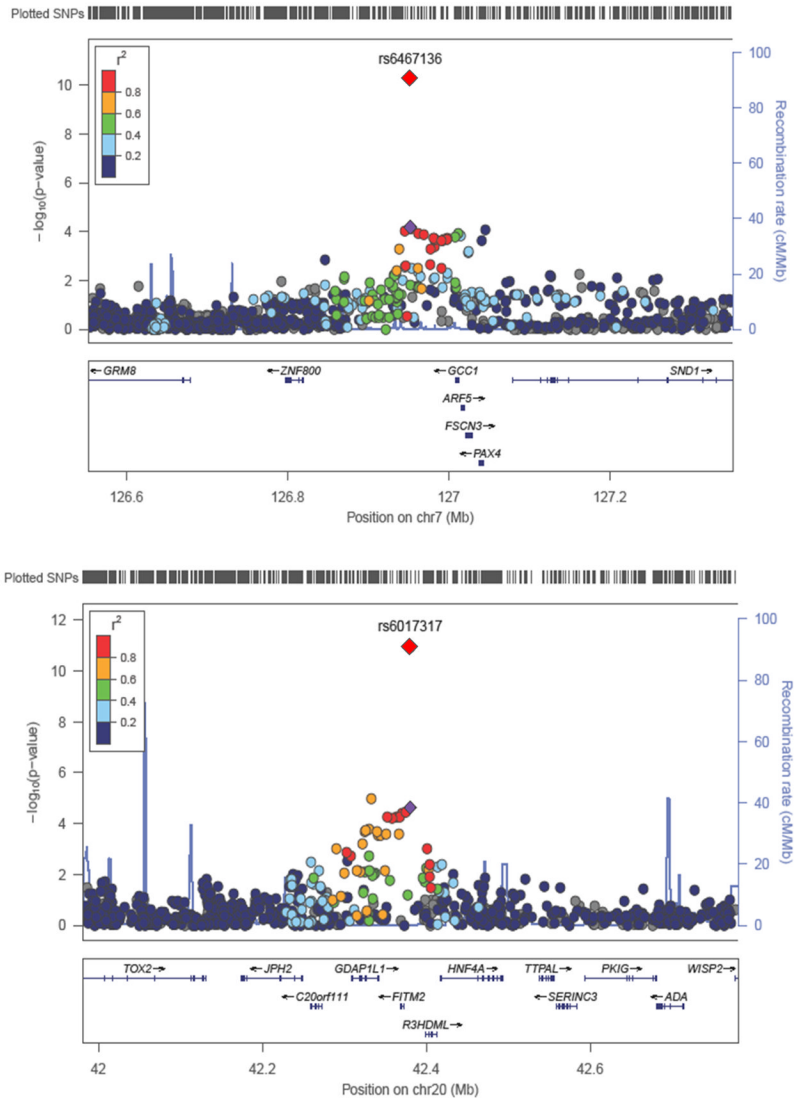
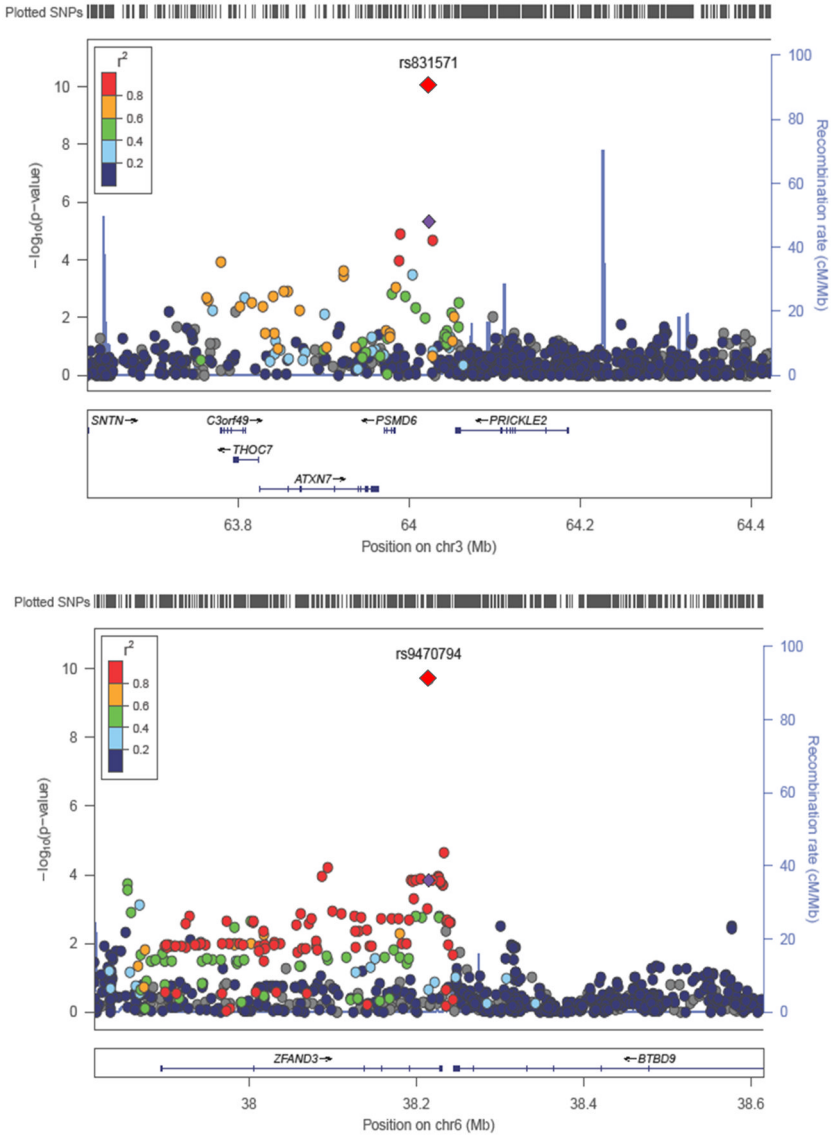


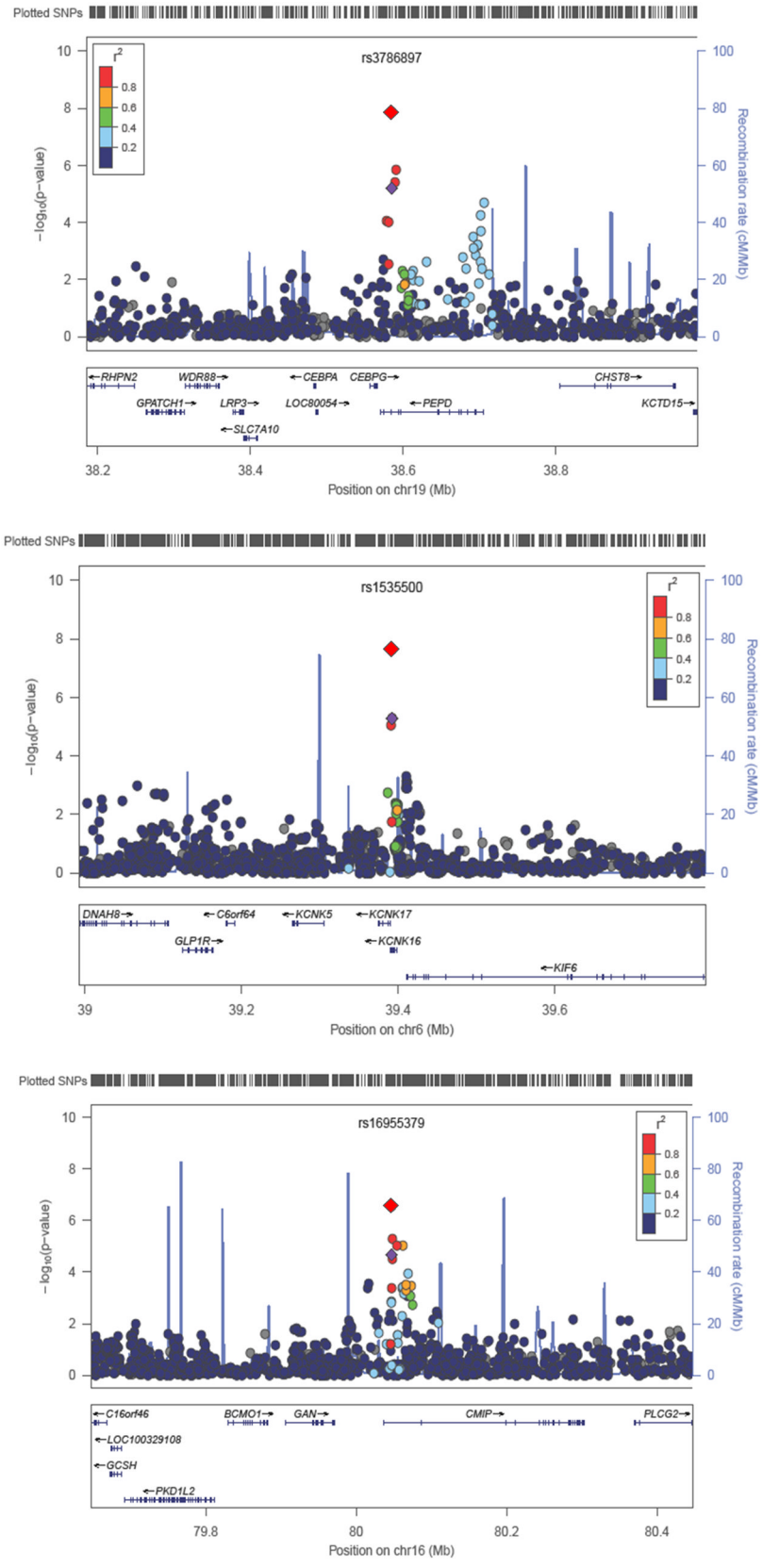
Figure 1. Genome-wide Manhattan plot for the EA T2D stage 1 meta-analysis

$-\log_{10}$(*P*-values) using the trend test are shown for SNPs distributed across the entire autosomal genome. Red dots at each locus indicate the signals with *P*-value $<10^{-6}$ detected in the genome-wide meta-analysis. A total of 1,934,619 SNPs that were present in ≥5 stage 1 studies were used to generate both plots.









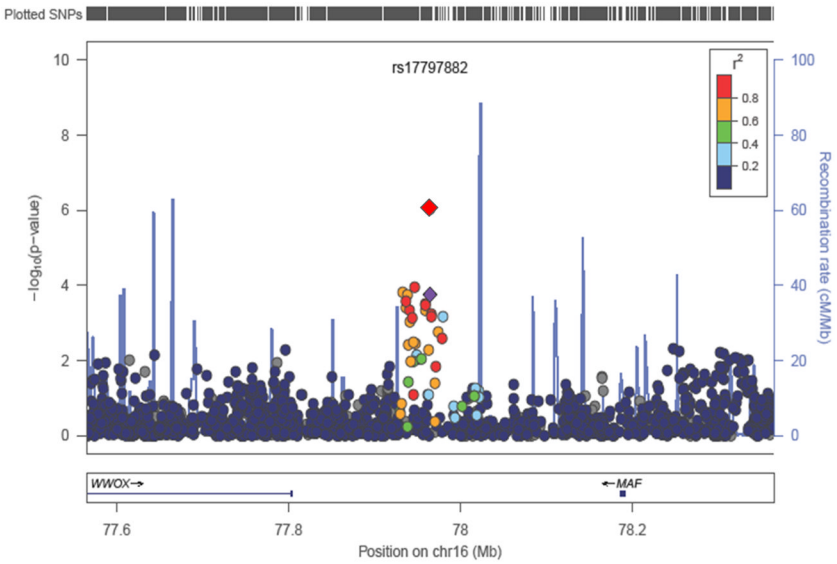


Figure 2. Regional association plots for new T2D loci

The upper panel indicates the positions of SNPs and the middle panel, shows regional association results from the genome-wide meta-analysis. The trend test $-\log_{10}$(*P*-values) are shown for SNPs distributed in a 0.8 Mb genomic region centered on the most strongly associated signal, which is depicted as a purple diamond in the stage 1 results and a red diamond for the combined stage 1+2+3 results. The lower panel shows the locations of known genes in the region. Genetic information is taken from Human Genome build hg18 and LD structure was based on 1000 Genomes (June 2010) HapMap JPT+CHB.

Table 1

Eight new T2D loci reaching genome-wide significance from combined meta-analysis of stage 1, 2 and 3.

SNP	Chr	Position (bp)	Nearby gene	Risk allele	Other allele	Stage 1 (discovery)		Stage 2 (*in silico* replication)		Stage3 (*de novo* replication)		Combined (stage 1+2+3)	
						OR (CI)	*P*-value	OR (CI)	*P*-value	OR (CI)	*P*-value	OR (CI)	*P*-value
						up to 6,952 cases and 11,865 controls		up to 5,843 cases and 4,574 controls		up to 12,284 cases and 13,172 controls		up to 25,079 cases and 29,611 controls	
Loci showing strong evidence of association with T2D													
rs6815464	4	1299901	*MAEA*	c	g	1.09 (1.04–1.14)	8.21E-04	1.13 (1.07–1.20)	3.67E-05	1.16 (1.11–1.20)	4.15E-15	1.13 (1.10–1.16)	1.57E-20
rs7041847	9	4277466	*GLIS3*	a	g	1.09 (1.04–1.14)	1.29E-04	1.09 (1.03–1.15)	2.20E-03	1.11 (1.07–1.15)	2.89E-09	1.10 (1.07–1.13)	1.99E-14
rs6017317	20	42380380	*FITM2-R3HDML-HNF4A*	g	t	1.10 (1.05–1.15)	2.43E-05	1.07 (0.99–1.15)	8.42E-02	1.10 (1.06–1.14)	3.96E-07	1.09 (1.07–1.12)	1.12E-11
rs6467136	7	126952194	*GCC1-PAX4*	g	a	1.12 (1.06–1.18)	6.47E-05	1.11 (1.04–1.18)	2.09E-03	1.10 (1.05–1.15)	2.31E-05	1.11 (1.07–1.14)	4.96E-11
rs831571	3	64023337	*PSMD6*	c	t	1.11 (1.06–1.17)	4.85E-06	1.06 (1.00–1.13)	4.46E-02	1.08 (1.05–1.12)	1.41E-05	1.09 (1.06–1.12)	8.41E-11
rs9470794	6	38214822	*ZFAND3*	c	t	1.11 (1.05–1.17)	1.45E-04	1.09 (1.02–1.17)	1.48E-02	1.16 (1.09–1.23)	3.20E-06	1.12 (1.08–1.16)	2.06E-10
rs3786897	19	38584848	*PEPD*	a	g	1.14 (1.08–1.20)	3.74E-06	1.05 (0.99–1.12)	1.28E-01	1.11 (1.04–1.17)	5.46E-04	1.10 (1.07–1.14)	1.30E-08
rs1535500	6	39392028	*KCNK16*	t	g	1.11 (1.06–1.16)	5.34E-06	1.07 (1.01–1.15)	3.33E-02	1.06 (1.02–1.10)	3.50E-03	1.08 (1.05–1.11)	2.30E-08
Loci showing moderate evidence of association with T2D													
rs16955379 *	16	80046874	*CMIP*	c	t	1.13 (1.07–1.20)	2.20E-05	1.10 (1.03–1.17)	6.59E-03	1.05 (1.01–1.10)	2.19E-02	1.08 (1.05–1.12)	2.84E-07
rs17797882	16	77964419	*WWOX*	t	c	1.12 (1.05–1.18)	1.76E-04	1.09 (1.02–1.16)	1.21E-02	1.06 (1.01–1.11)	1.61E-02	1.08 (1.05–1.12)	9.49E-07

* Proxy SNP, rs9930117 (r^2 =1), was genotyped in the stage3 CAGE study.