nature genetics

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

Yoon Shin Cho[1,46], Chien-Hsiun Chen[2,3,46], Cheng Hu[4,46], Jirong Long[5,46], Rick Twee Hee Ong[6,46], Xueling Sim[7,46], Fumihiko Takeuchi[8,46], Ying Wu[9,46], Min Jin Go[1,46], Toshimasa Yamauchi[10,46], Yi-Cheng Chang[11,46], Soo Heon Kwak[12,46], Ronald C W Ma[13,46], Ken Yamamoto[14,46], 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[37,47], Kyong Soo Park[12,38,47], Weiping Jia[4,47], Lee-Ming Chuang[11,39,47], Juliana C N Chan[13,47], Shiro Maeda[39,47], Takashi Kadowaki[10,47], Jong-Young Lee[1,47], Jer-Yuarn Wu[2,3,47], Yik Ying Teo[6,7,22,25,41,47], E Shyong Tai[22,42,43,47], Xiao Ou Shu[5,47], Karen L Mohlke[9,47], Norihiro Kato[8,47], Bok-Ghee Han[1,47] & Mark Seielstad[25,44,45,47]

We conducted a three-stage genetic study to identify susceptibility loci for type 2 diabetes (T2D) in east Asian populations. We followed our stage 1 meta-analysis of eight T2D genome-wide association studies (6,952 cases with T2D and 11,865 controls) with a stage 2 *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 mapped in or near *GLIS3*, *PEPD*, *FITM2-R3HDML-HNF4A*, *KCNK16*, *MAEA*, *GCC1-PAX4*, *PSMD6* and *ZFAND3*. *GLIS3*, which is 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 an association with T2D for *PEPD*[5] and *HNF4A*[6,7] has been shown in previous studies. *KCNK16* may regulate glucose-dependent insulin secretion in the pancreas. These findings, derived from an east Asian population, provide new perspectives on the etiology of T2D.

T2D is a major public health problem with a rapidly rising global prevalence[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 has considerable heritability. Within only the last 3 years, genetic studies have produced a rapidly lengthening list of loci harboring disease-predisposing variations[10]. To date, genetic variants at 45 loci have been identified for T2D[10,11]. Despite these advances toward a better understanding of the genetic basis of T2D, its heritability has not been fully explained[12]. In addition, most of the T2D loci were detected initially in population samples of European origin, with the exceptions of *KCNQ1*, *UBE2E2* and *C2CD4A-C2CD4B*, which were first identified in studies of east Asian populations[13–15]. Additional efforts involving east Asian populations identified variants associated with T2D at the *SPRY2*, *PTPRD* and *SRR* loci[5,16,17]. However, these associations need more validation from additional studies of east Asians as well as in studies of other populations. A large meta-analysis in east Asians would be expected to identify new 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 in these populations, T2D epidemiology also differs considerably between European populations and east Asian populations. In east Asians, the rates of diabetes are often higher at lower average body mass indices (BMIs)[18], suggesting that some different pathways may be involved in pathogenesis of T2D in east Asians and Europeans.

To discover new T2D loci, we conducted a three-stage association study in individuals of east Asian descent (**Supplementary Fig. 1**). We performed the stage 1 meta-analysis by combining eight T2D genome-wide association studies (GWAS) participating in the Asian Genetic Epidemiology Network (AGEN) consortium (6,952 cases and 11,865 controls) with association data for 2,626,356 imputed and genotyped autosomal SNPs, and we used the inverse-variance method

A full list of affiliations are at the end of article.

Received 12 April; accepted 2 November; published online 11 December 2011; doi:10.1038/ng.1019

for fixed effects for the statistical analyses (**Supplementary Table 1**). All imputed and genotyped SNPs (minor allele frequency (MAF) > 0.01) passed quality control filters in each of the eight stage 1 datasets prior to conducting the meta-analysis (**Supplementary Table 2**). The genomic control inflation factor (λ) for the meta-analysis was 1.046 (and was less than 1.062 for each of the individual studies), indicating that the results seen in stage 1 were probably not the result of population stratification (**Supplementary Fig. 2**). Individuals from each component study that participated in stage 1 mainly clustered together with the samples from the CHB/JPT HapMap population in the principal component analysis plot (**Supplementary Fig. 3**), further showing the similarity in ethnicity between the stage 1 samples. Most signals showing strong evidence for T2D associations were in known T2D genes (**Fig. 1**). Stage 1 *P* values, odds ratios (ORs) and average risk allele frequencies for 45 previously reported T2D-associated SNPs are listed in **Supplementary Table 3**.

After removing known T2D variants, we selected 297 SNPs from independent loci from the stage 1 meta-analysis based on our arbitrary inclusion criteria for *in silico* follow-up replication: meta-analysis $P < 5 \times 10^{-4}$ (based on the divergence between the observed and expected *P* values on the quantile-quantile plot; **Supplementary Fig. 2**), heterogeneity $P > 0.01$ and at least seven studies having been included in the meta-analysis (**Supplementary Table 4**). We took a total of 3,756 SNPs, including the 297 selected SNPs and their proxies ($r^2 > 0.8$ based on phase 2 CHB/JPT HapMap data), forward to stage 2 (*in silico* replication) in three independent GWAS (5,843 cases and 4,574 controls). After a meta-analysis that combined stage 1 and 2 data for 3,756 SNPs, we selected the 19 SNPs that showed the most compelling evidence for association (stage 1 and 2 combined $P < 10^{-5}$) (**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** and **2**). This resulted in eight new T2D loci that reached genome-wide significance in the combined meta-analysis across all three stages (**Table 1** and **Fig. 2**).

Three of these T2D-associated loci were previously associated with metabolic traits or related diseases or were suggestively associated with T2D. We detected one such locus 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 key player in the regulation of pancreatic beta cell development and insulin gene expression[1,2]. SNPs in high linkage disequilibrium (LD) with this locus have been implicated in association with type 1 diabetes (T1D)[19] and fasting plasma glucose[3]. The second such locus, on 19q13, is located in an intron of *PEPD*. Several SNPs (lead SNP: rs10425678) in this gene were previously associated with T2D in a Japanese population[5]. However, the SNP in *PEPD* identified in our study (rs3786897) is not in LD with those identified in the Japanese population ($r^2 = 0.008$, $D' = 0.143$ between rs3786897 and rs10425678 based on phase 2 CHB/JPT HapMap data), and our GWAS 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], and the function of *R3HDML* is not known. Mutations in *HNF4A* cause maturity onset diabetes of the young type 1 (ref. 21). Common variants in the P2 promoter region of this gene (rs1884613 and rs2144908) have been associated with T2D in a population-specific manner[6,22]. The SNP near *FITM2-R3HDML-HNF4A* identified in our study (rs6017317) is not in strong LD with the *HNF4A* P2 promoter SNPs ($r^2 = 0.23$–0.25, $D' = 0.50$–0.54 between rs6017317 and rs1884613 or rs2144908 based on phase 2 CHB/JPT HapMap data), indicating that rs6017317 is a new T2D signal in the 20q13.12 region where *HNF4A* resides.

The other 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*, which is 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 the signal reaching genome-wide significance in *KCNK16* (rs1535500) is rs11756091 ($r^2 = 0.977$, $D' = 1.0$ based on phase 2 CHB/JPT HapMap data), which encodes a substitution of proline to histidine in two isoforms of KCNK16. This variant or others influencing *KCNK16* may result in the defective regulation of potassium channel activity that contributes to the etiology of T2D[24]. *MAEA* encodes a protein that has a role in erythroblast enucleation and in the development of mature macrophages[25]. A gene-set analysis of the stage 1 *P* values using GSA-SNP[26] indicated that *MAEA* belongs to a group of genes that previously showed significant association with T2D and includes *IDE*, which is located at a known T2D susceptibility locus[27] (stage 1 $P = 1.41 \times 10^{-7}$ for rs6583826 at the *IDE* locus in this study). The GRIP-domain–containing protein that is encoded by *GCC1* might have a role in the organization of the trans-Golgi network, which is involved in membrane transport[28]. *PAX4*, which is only 30 kb away from *GCC1*, is a good candidate for T2D given its involvement in pancreatic islet development. *PAX4* was recently implicated in a Japanese individual with maturity onset diabetes of the young[29]. The expression product of *PSMD6*, which acts as a regulatory subunit of the 26S proteasome, is probably involved in the ATP-dependent degradation of ubiquitinated proteins[30]. Although the function of *ZFAND3* has not been fully elucidated, it is noteworthy that a member of the same gene family, *ZFAND6*, is present along with *FAH* at a previously detected T2D locus[31]. We examined whether eight new loci are potentially associated with T2D through an effect on obesity, as is the case with *FTO*[32]. All of the T2D association signals we initially detected remained after adjustment for BMI (**Supplementary Table 6**), indicating that the associations with T2D of these eight loci are not mediated through an effect on obesity.

In addition to the eight loci reaching genome-wide significance, we identified two loci showing moderate evidence (combined $P < 10^{-6}$) of association with T2D, including *WWOX* and *CMIP* loci (**Table 1**). We obtained the association results for these ten loci in GWAS data from up to 47,117 European samples generated by the DIAGRAM consortium (DIAGRAM+ is the current version of dataset)[31].

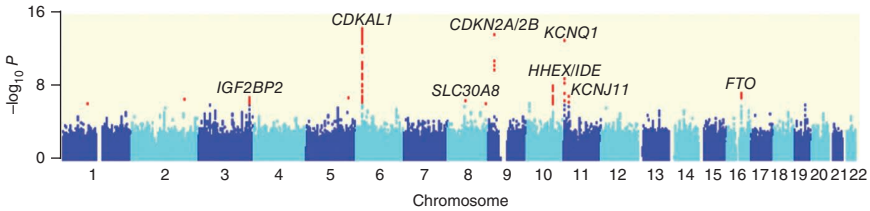

Figure 1 Genome-wide Manhattan plot for the east Asian T2D stage 1 meta-analysis. Shown are the $-\log_{10} P$ values using the trend test for SNPs distributed across the entire autosomal genome. The red dots at each locus indicate the signals with $P < 10^{-6}$ detected in the genome-wide meta-analysis. A total of 1,934,619 SNPs that were present in at least five stage 1 studies were used to generate the plot.

Table 1 Eight new T2D loci reaching genome-wide significance from a combined meta-analysis of stages 1, 2 and 3

SNP	Chr.	Position (bp)	Nearby gene	Risk allele	Other allele	Stage 1 (discovery)[a]		Stage 2 (*in silico* replication)[b]		Stage 3 (*de novo* replication)[c]		Combined (stages 1, 2 and 3)[d]	
						OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Loci showing strong evidence of association with T2D													
rs6815464	4	1,299,901	*MAEA*	C	G	1.09 (1.04–1.14)	8.21×10^{-4}	1.13 (1.07–1.20)	3.67×10^{-5}	1.16 (1.11–1.20)	4.15×10^{-15}	1.13 (1.10–1.16)	1.57×10^{-20}
rs7041847	9	4,277,466	*GLIS3*	A	G	1.09 (1.04–1.14)	1.29×10^{-4}	1.09 (1.03–1.15)	2.20×10^{-3}	1.11 (1.07–1.15)	2.89×10^{-9}	1.10 (1.07–1.13)	1.99×10^{-14}
rs6017317	20	42,380,380	*FITM2-R3HDML-HNF4A*	G	T	1.10 (1.05–1.15)	2.43×10^{-5}	1.07 (0.99–1.15)	8.42×10^{-2}	1.10 (1.06–1.14)	3.96×10^{-7}	1.09 (1.07–1.12)	1.12×10^{-11}
rs6467136	7	126,952,194	*GCC1-PAX4*	G	A	1.12 (1.06–1.18)	6.47×10^{-5}	1.11 (1.04–1.18)	2.09×10^{-3}	1.10 (1.05–1.15)	2.31×10^{-5}	1.11 (1.07–1.14)	4.96×10^{-11}
rs831571	3	64,023,337	*PSMD6*	C	T	1.11 (1.06–1.17)	4.85×10^{-6}	1.06 (1.00–1.13)	4.46×10^{-2}	1.08 (1.05–1.12)	1.41×10^{-5}	1.09 (1.06–1.12)	8.41×10^{-11}
rs9470794	6	38,214,822	*ZFAND3*	C	T	1.11 (1.05–1.17)	1.45×10^{-4}	1.09 (1.02–1.17)	1.48×10^{-2}	1.16 (1.09–1.23)	3.20×10^{-6}	1.12 (1.08–1.16)	2.06×10^{-10}
rs3786897	19	38,584,848	*PEPD*	A	G	1.14 (1.08–1.20)	3.74×10^{-6}	1.05 (0.99–1.12)	1.28×10^{-1}	1.11 (1.04–1.17)	5.46×10^{-4}	1.10 (1.07–1.14)	1.30×10^{-8}
rs1535500	6	39,392,028	*KCNK16*	T	G	1.11 (1.06–1.16)	5.34×10^{-6}	1.07 (1.01–1.15)	3.33×10^{-2}	1.06 (1.02–1.10)	3.50×10^{-3}	1.08 (1.05–1.11)	2.30×10^{-8}
Loci showing moderate evidence of association with T2D													
rs16955379[e]	16	80,046,874	*CMIP*	C	T	1.13 (1.07–1.20)	2.20×10^{-5}	1.10 (1.03–1.17)	6.59×10^{-3}	1.05 (1.01–1.10)	2.19×10^{-2}	1.08 (1.05–1.12)	2.84×10^{-7}
rs17797882	16	77,964,419	*WWOX*	T	C	1.12 (1.05–1.18)	1.76×10^{-4}	1.09 (1.02–1.16)	1.21×10^{-2}	1.06 (1.01–1.11)	1.61×10^{-2}	1.08 (1.05–1.12)	9.49×10^{-7}

[a]Up to 6,952 cases and 11,865 controls. [b]Up to 5,843 cases and 4,574 controls. [c]Up to 12,284 cases and 13,172 controls. [d]Up to 25,079 cases and 29,611 controls. [e]The proxy SNP rs9930117 ($r^2 = 1$) was genotyped in the stage 3 CAGE study.

The DIAGRAM-generated results for these loci indicated that three loci, including the *FITM2-R3HDML-HNF4A* (rs6017317: $P = 1.47 \times 10^{-2}$, OR = 1.07), *CMIP* (rs16955379: $P = 3.33 \times 10^{-2}$, OR = 1.20) and *MAEA* (using the proxy SNP for rs6815464, rs11247991 ($r^2 = 0.96$): $P = 6.56 \times 10^{-3}$, OR = 1.19) loci, were modestly associated with T2D, whereas a locus in *GLIS3* (rs7041847: $P = 6.43 \times 10^{-2}$, OR = 1.04) was nominally associated with T2D. The direction of effect was consistent in four (*PSMD6*, *PEPD*, *WWOX* and *KCNK16*) of the six loci that were not replicated in DIAGRAM+ (**Supplementary Table 5**).

We analyzed the functional connections among the 10 new T2D genes and the 28 known T2D genes that we replicated in this study (**Supplementary Table 3**) using GRAIL[33]. The connection results highlighted notable biological functions for sets of genes within T2D-associated regions (**Supplementary Fig. 4** and **Supplementary Tables 7** and **8**). For example, *KCNK16* has strong connections with previously known T2D genes encoding potassium channels (*KCNJ11* and *KCNQ1*), implying that is has a physiological role in the regulation of potassium transport in pancreatic cells.

We examined the association between each new T2D SNP and the expression of genes within 1 Mb of these SNPs by an expression quantitative trait locus (eQTL) analysis using the data from the MuTHER consortium. One SNP (rs3786897) in an intron of *PEPD* was highly associated with the mRNA expression of *PEPD* in the adipose tissue of 776 individuals of European ancestry ($P_{eQTL} = 2.14 \times 10^{-8}$) (**Supplementary Table 9**). However, this SNP did not show an association with T2D in populations of European ancestry, thus the importance of this finding is unclear.

We considered the possibility that autoimmune diabetes (rather than T2D) may be driving some of the signals that we observed. First, the cases from all the studies we examined predominantly had adult-onset diabetes (age of disease onset ≥30 years), and none of the clinically diagnosed subjects had T1D, which is defined by the presence of acute ketosis and the continuous requirement of insulin beginning within 1 year after diagnosis. Second, we researched the associations for all known T1D-associated variants in our dataset. Only a small number of loci showed association after this analysis (**Supplementary Table 10**). These results are in distinct contrast to those for known T2D-associated variants, many of which replicated in our study (**Supplementary Table 3**), further suggesting that our findings are most relevant to T2D. Third, as 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 individulas with positive glutamic acid decarboxylase (GAD) antibodies had been excluded (as individuals with T1D are positive for GAD antibodies, whereas individuals with T2D are not) (data not shown). In each study, the associations between this SNP and diabetes were the same as the association found when we included all the samples (meta-analysis $P = 3.4 \times 10^{-4}$, 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 in healthy children and adolescents[4], is consistent with the hypothesis that SNPs at this locus may affect fasting glucose homeostasis rather than the immune system. Taken together, it is unlikely that a substantial proportion of the positive associations observed in our study were driven by autoimmune diabetes.

This study is the largest GWAS meta-analysis, to our knowledge, 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 seems to be highly specific to east Asian populations (**Supplementary Table 5**), whereas the association of

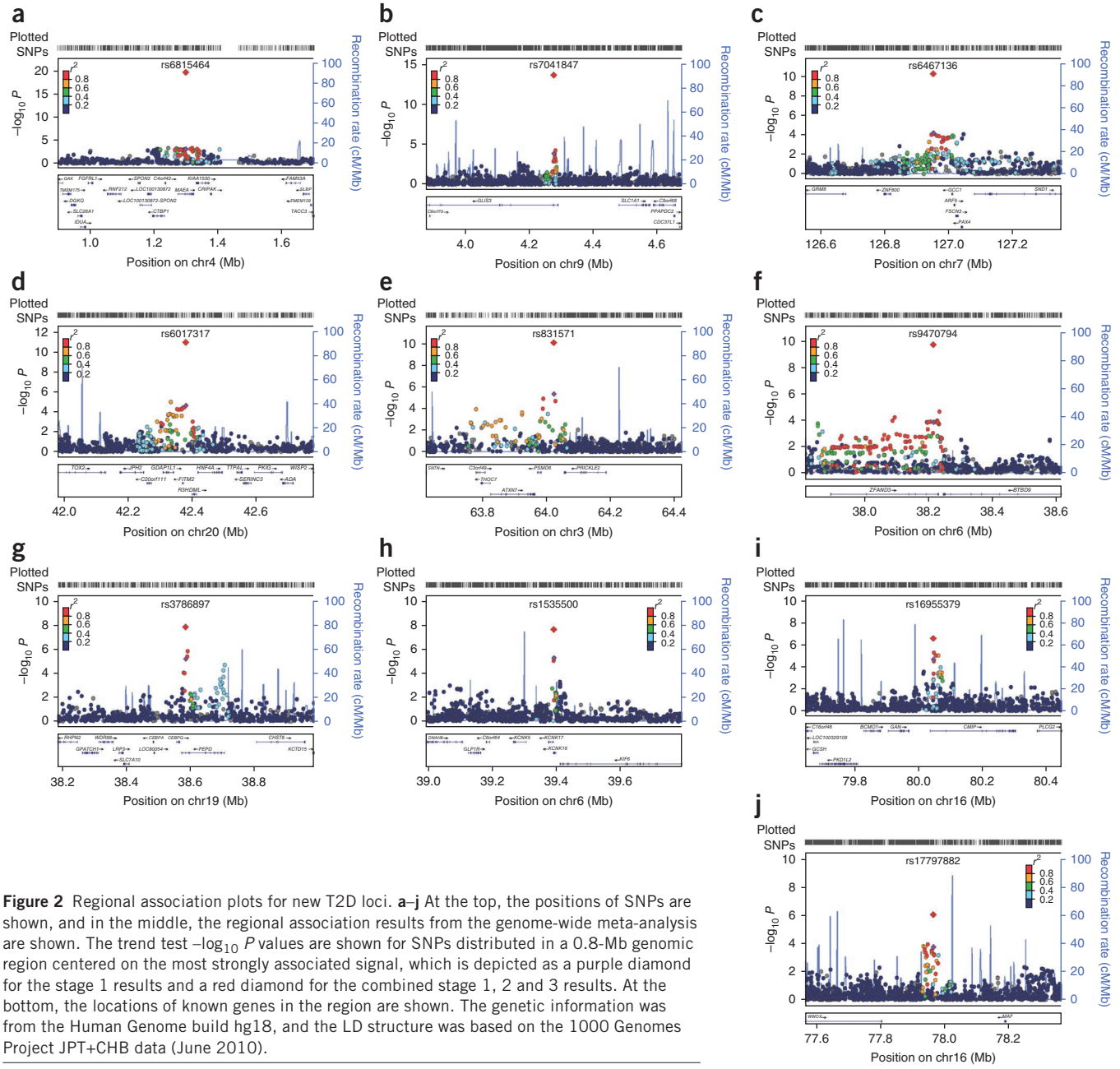

Figure 2 Regional association plots for new T2D loci. **a**–**j** At the top, the positions of SNPs are shown, and in the middle, the regional association results from the genome-wide meta-analysis are shown. The trend test $-\log_{10} P$ values are shown for SNPs distributed in a 0.8-Mb genomic region centered on the most strongly associated signal, which is depicted as a purple diamond for the stage 1 results and a red diamond for the combined stage 1, 2 and 3 results. At the bottom, the locations of known genes in the region are shown. The genetic information was from the Human Genome build hg18, and the LD structure was based on the 1000 Genomes Project JPT+CHB data (June 2010).

rs11634397 near *ZFAND6* seems to be specific to European populations (**Supplementary Table 3**). We observed a substantial difference in the risk allele frequencies of both loci between the two continental (Asian and European) populations (rs9470794: risk allele frequency (RAF) = 0.32 for the Asian CHB/JPT HapMap population compared to RAF = 0.12 for the European CEU HapMap population; rs11634397: RAF = 0.07 for CHB/JPT compared to RAF = 0.64 for CEU). Although these loci are related to T2D differently in the two populations (the *ZFAND3* locus is specific to Asians, whereas the *ZFAND6* locus to Europeans), these results lead to speculation that the broader A20 domain-containing zinc finger protein family has a role in the etiology of T2D. Additional population-specific T2D loci were also suggested by our analysis, for example, *WWOX* (rs17797882) (**Supplementary Table 5**) in east Asians and *ZBED3* (rs4457053) (**Supplementary Table 3**) in Europeans. Despite the lack of clear physiological evidence on T2D pathogenesis, these findings may provide clues to help understand T2D phenotypes characteristic of each population, for example, the high rates of diabetes seen at lower average BMIs in east Asians.

URLs. IMPUTE, http://mathgen.stats.ox.ac.uk/impute/impute.html; MACH, http://www.sph.umich.edu/csg/abecasis/MACH/; BEAGLE, http://faculty.washington.edu/browning/beagle/beagle.html; METAL, http://www.sph.umich.edu/csg/abecasis/Metal; WGAViewer, http://compute1.lsrc.duke.edu/softwares/WGAViewer/; SNAP, http://www.broadinstitute.org/mpg/snap/; LocusZoom, http://csg.sph.umich.edu/locuszoom/; GenABEL, http://www.genabel.org/; ProbABEL, http://www.genabel.org/packages/ProbABEL.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

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.

The Japan Cardiometabolic Genome Epidemiology (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; the Grant of National Center for Global Health and Medicine (NCGM).

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 US National Institutes of Health grants DK078150, TW05596, HL085144 and TW008288 and pilot funds from grants RR20649, ES10126, and DK56350.

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 Special Administrative Region (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.

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 (2011-N73005-00), the Republic of Korea.

The work by the National Taiwan University Hospital 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.

The work by the Shanghai Diabetes Genetic Study was supported in part by the US National Institutes of Health grants R01CA124558, R01CA64277, R01CA70867, R01CA90899, R01CA100374, R01CA118229, R01CA92585, UL1 RR024975, DK58845 and HG004399, the Department of Defense Idea Award BC050791, Vanderbilt Ingram professorship funds and the Allen Foundation Fund. We thank the dedicated investigators and staff members from research teams at Vanderbilt University, Shanghai Cancer Institute and the Shanghai Institute of Preventive Medicine, and especially, the study participants for their contributions in the studies.

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

The work of the Shanghai Jiao Tong University Diabetes Study was supported by grants from the National 863 Program (2006AA02A409) and the major program of the Shanghai Municipality for Basic Research (08dj1400601), China.

The work of the Seoul National University Hospital was supported by grants from the Korea Health 21 R&D Project, Ministry of Health & Welfare (00-PJ3-PG6-GN07-001) and the World Class University project of the Ministry of Education, Science and Technology (MEST) and National Research Foundation (NRF) (R31-2008-000-10103-0), Korea. The Singapore Prospective Study Program (SP2) was funded through grants from the Biomedical Research Council of Singapore (BMRC05/1/36/19/413 and 03/1/27/18/216) and the National Medical Research Council of Singapore (NMRC/1174/2008). E.S.T. also receives additional support from the National Medical Research Council through a clinicians scientist award (NMRC/CSA/008/2009). The Singapore Malay Eye Study (SiMES) was funded by the National Medical Research Council (NMRC0796/2003 and NMRC/STaR/0003/2008) and Biomedical Research Council (BMRC, 09/1/35/19/616). Y.Y.T. acknowledges support from the Singapore National Research Foundation, NRF-RF-2010-05. The Genome Institute of Singapore carried out all the genotyping for the samples from Singapore also provided funding for the genotyping of the samples from SP2.

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.N.C. The experiments were conceived of 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.B.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.C.W.M. The experiments were performed by J.L., M.S., J.J.L., J.-Y.W., S.M., R.Z., K.Y., Y.-C.C., T.-J.C., L.-M.C. and S.H.K. Statistical analyses 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.C., S.H.K., A.P.M. and R.C.W.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.C., S.H.K., Y.S.C., S.K., Å.K.H. and R.C.W.M. The 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.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.N.C. The manuscript was written by Y.S.C., M.S. and E.S.T. All authors reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Kang, H.S. *et al.* Transcription factor Glis3, a novel critical player in the regulation of pancreatic beta-cell development and insulin gene expression. *Mol. Cell. Biol.* **29**, 6366–6379 (2009).
2. Yang, Y., Chang, B.H., Samson, S.L., Li, M.V. & Chan, L. The Kruppel-like zinc finger protein Glis3 directly and indirectly activates insulin gene transcription. *Nucleic Acids Res.* **37**, 2529–2538 (2009).
3. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
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* **60**, 1805–1812 (2011).
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* **58**, 1690–1699 (2009).
6. Barroso, I. *et al.* Population-specific risk of type 2 diabetes conferred by HNF4A P2 promoter variants: a lesson for replication studies. *Diabetes* **57**, 3161–3165 (2008).
7. Silander, K. *et al.* Genetic variation near the hepatocyte nuclear factor-4 α gene predicts susceptibility to type 2 diabetes. *Diabetes* **53**, 1141–1149 (2004).
8. Zimmet, P., Alberti, K.G. & Shaw, J. Global and societal implications of the diabetes epidemic. *Nature* **414**, 782–787 (2001).
9. Tkác, I. Metabolic syndrome in relationship to type 2 diabetes and atherosclerosis. *Diabetes Res. Clin. Pract.* **68** (suppl. 1), S2–S9 (2005).
10. Prokopenko, I., McCarthy, M.I. & Lindgren, C.M. Type 2 diabetes: new genes, new understanding. *Trends Genet.* **24**, 613–621 (2008).
11. Rung, J. *et al.* Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat. Genet.* **41**, 1110–1115 (2009).
12. Manolio, T.A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
13. Yasuda, K. *et al.* Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nat. Genet.* **40**, 1092–1097 (2008).
14. Unoki, H. *et al.* SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in east Asian and European populations. *Nat. Genet.* **40**, 1098–1102 (2008).
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**, 864–868 (2010).
16. Tsai, F.J. *et al.* A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. *PLoS Genet.* **6**, e1000847 (2010).
17. Shu, X.O. *et al.* Identification of new genetic risk variants for type 2 diabetes. *PLoS Genet.* **6**, e1001127 (2010).
18. Stommel, M. & Schoenborn, C.A. Variations in BMI and prevalence of health risks in diverse racial and ethnic populations. *Obesity (Silver Spring)* **18**, 1821–1826 (2010).
19. Barrett, J.C. *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41**, 703–707 (2009).
20. Kadereit, B. *et al.* Evolutionarily conserved gene family important for fat storage. *Proc. Natl. Acad. Sci. USA* **105**, 94–99 (2008).
21. Nakajima, H. *et al.* Hepatocyte nuclear factor-4 α gene mutations in Japanese non-insulin dependent diabetes mellitus (NIDDM) patients. *Res. Commun. Mol. Pathol. Pharmacol.* **94**, 327–330 (1996).
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 α (HNF4A) P2 region are associated with type 2 diabetes in Scandinavians. *Diabetes* **56**, 3112–3117 (2007).
23. Girard, C. *et al.* Genomic and functional characteristics of novel human pancreatic 2P domain K^+ channels. *Biochem. Biophys. Res. Commun.* **282**, 249–256 (2001).
24. Ashcroft, F.M. ATP-sensitive potassium channelopathies: focus on insulin secretion. *J. Clin. Invest.* **115**, 2047–2058 (2005).
25. Soni, S. *et al.* Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. *J. Biol. Chem.* **281**, 20181–20189 (2006).
26. Nam, D., Kim, J., Kim, S.Y. & Kim, S. GSA-SNP: a general approach for gene set analysis of polymorphisms. *Nucleic Acids Res.* **38**, W749–W754 (2010).
27. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).

28. Luke, M.R., Houghton, F., Perugini, M.A. & Gleeson, P.A. The trans-Golgi network GRIP-domain proteins form α-helical homodimers. *Biochem. J.* **388**, 835–841 (2005).
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.* **223**, 113–118 (2011).
30. Wang, X. *et al.* Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry* **46**, 3553–3565 (2007).
31. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).
32. Frayling, T.M. *et al.* A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 (2007).
33. Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* **5**, e1000534 (2009).

[1]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Cheongwon-gun, Gangoe-myeon, Yeonje-ri, Korea. [2]Institute of Biomedical Sciences, Academia Sinica, Nankang, Taipei, Taiwan. [3]School of Chinese Medicine, China Medical University, 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, Shanghai, China. [5]Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [6]Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore. [7]Centre for Molecular Epidemiology, National University of Singapore, Singapore, Singapore. [8]Research Institute, National Center for Global Health and Medicine, Shinjuku-ku, Tokyo, Japan. [9]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA. [10]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [11]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan. [12]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea. [13]Department of Medicine and Therapeutics, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China. [14]Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka, Japan. [15]Department of Nutrition, University of North Carolina, Chapel Hill, North Carolina, USA. [16]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore. [17]Department of Ophthalmology, National University of Singapore, Singapore, Singapore. [18]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China. [19]Department of Nutrition and Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [20]Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Korea. [21]School of Systems Biomedical Science, Soongsil University, Dongjak-gu, Seoul, Korea. [22]Department of Epidemiology and Public Health, National University of Singapore, Singapore, 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, North Carolina, USA. [25]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore. [26]Shanghai Institute of Preventive Medicine, Shanghai, China. [27]Division of Endocrinology and Diabetes, Department of Internal Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan. [28]Department of Diabetes and Endocrinology, Chubu Rosai Hospital, Nagoya, Japan. [29]Centre for Eye Research Australia, University of Melbourne, East Melbourne, Victoria, Australia. [30]Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya, Japan. [31]Department of Geriatric Medicine, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan. [32]Department of Endocrinology, Metabolism and Diabetes, Kinki University School of Medicine, Osaka-sayama, Osaka, Japan. [33]Department of Preventive Medicine, Ajou University School of Medicine, Suwon, 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. [36]A full list of members is provided in the **Supplementary Note**. [37]Department of Internal Medicine and Bioregulatory Science, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan. [38]World Class University program, Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology and College of Medicine, Seoul National University, Seoul, Korea. [39]Graduate Institute of Clinical Medicine, National Taiwan University School of Medicine, Taipei, Taiwan. [40]Laboratory for Endocrinology and Metabolism, RIKEN Center for Genomic Medicine, Yokohama, Japan. [41]Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore. [42]Department of Medicine, National University of Singapore, Singapore, Singapore. [43]Duke-National University of Singapore Graduate Medical School, Singapore, Singapore. [44]Institute for Human Genetics, University of California, San Francisco, California, USA. [45]Blood Systems Research Institute, San Francisco, California, USA. [46]These authors contributed equally to this work. [47]These authors jointly directed this work. 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).

ONLINE METHODS

Study subjects. Stage 1 subjects were drawn from eight T2D GWAS participating in the AGEN consortium, which was organized to enable genetic studies on diverse complex traits in 2010. These eight studies included 6,952 cases with T2D 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). Subjects in stage 2 included 5,843 cases with T2D and 4,574 controls from three independent GWAS, 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 with T2D and 13,172 controls from five different studies, 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 used in each study included in stages 1, 2, and 3 are described in **Supplementary Table 1** and the **Supplementary Note**. Each study obtained approval from the appropriate institutional review boards of each participating institution, and written informed consent was obtained from all participants. The three-stage design of the overall study is depicted in **Supplementary Figure 1**.

Genotyping and imputation. Subjects for the stage 1 and 2 analyses were genotyped with high-density SNP typing platforms that covered the entire human genome. In most of the studies, only unrelated samples with missing genotype call rates below 5% were included for subsequent GWAS analyses. For the genome-wide association meta-analysis, each study participating in stages 1 and 2 performed SNP imputation. IMPUTE, MACH or BEAGLE (see URLs) were used, together with haplotype reference panels from the JPT and CHB samples that are available in the HapMap database (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 the association analysis. Genotyping for the stage 3 analysis was carried out using TaqMan, Sequenom MassARRAY or the Beckman SNP Stream method. All SNPs included in stage 3 had a genotype success rate of >98% (**Supplementary Table 2**).

Statistical analyses, analysis tools and SNP prioritization for stages 2 and 3. Associations between SNPs and T2D were tested by logistic regression with an additive model (1 degree of freedom) after adjustment for sex. Other adjustments were permitted according to the situations in the individual studies. The meta-analysis was performed using an inverse-variance method assuming fixed effects, with a Cochran's *Q* test to assess between-study heterogeneity. METAL software (see URLs) was used for all meta-analyses. A plot of the negative log of the association results from the stage 1 meta-analysis, by chromosome, was generated using WGAViewer software (see URLs).The quantile-quantile plot was constructed by plotting the distribution of observed *P* values for the given SNPs against the theoretical distribution of the expected *P* values for T2D[34]. The genomic control inflation factor, λ, was calculated by dividing the median χ^2 statistics by 0.456 (ref. 35) for individual GWAS, as well as for the stage 1 GWAS meta-analysis. We did not correct for genomic control in the stage 1 analyses because the inflation was modest, suggesting that population structure is unlikely to cause substantial inflation of the stage 1 results (**Supplementary Table 2**). The selection criteria for the lead SNPs to take forward to stage 2 *in silico* replication analysis were as follows: (i) stage 1 meta-analysis $P < 5 \times 10^{-4}$ (based on the divergence between the observed and expected *P* values on the quantile-quantile plot; **Supplementary Fig. 2**); (ii) heterogeneity $P > 0.01$; and (iii) at least seven studies having been included in the stage 1 meta-analysis (**Supplementary Table 4**). After removing known variants associated with T2D, proxies for each lead SNP ($r^2 > 0.8$) were selected using the SNAP software (see URLs). The replication genotyping for stage 3 was performed for the new SNPs having a stage 2 combined $P < 10^{-5}$. Regional association results from genome-wide meta-analysis were plotted using LocusZoom software (see URLs) for SNPs reaching genome-wide significance from the combined meta-analysis of stages 1, 2 and 3.

Principal components analysis. A list of 76,534 common SNPs across the Illumina 550, 610 and 1M and Affymetrix 5.0 and 6.0 arrays were first selected. This set of SNPs in the Asian (CHB+JPT) HapMap II samples was then trained to generate a list of 44,524 SNPs having pairwise LD < 0.3 in a sliding window of 50 SNPs. Individuals from each component study and from HapMap II were plotted based on the first two eigenvectors produced by the principal components analysis.

eQTL analysis. Gene expression information from 776 adipose tissues, 667 skin tissues and 777 lymphoblastoid cell lines was obtained from the MuTHER consortium[36]. The eQTL data for eight of the ten T2D loci identified in this study were available in the MuTHER dataset. Most of those loci passed the filtering criteria, such as MAF > 5% and INFO > 0.8, except for rs16955379, which has MAF = 1.5% in the MuTHER data set. Two of the ten loci that were used in the eQTL analysis, rs6815464 (on chromosome 4) and rs17797882 (on chromosome 16), are not included in the MuTHER data set. Association between each SNP with a significant association to T2D and the normalized mRNA expression values of genes within 1 Mb of the lead SNP were performed with the GenABEL and ProbABEL package (see URLs) 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 $P = 5.06 \times 10^{-5}$ in adipose tissue, $P = 3.81 \times 10^{-5}$ in skin and $P = 7.80 \times 10^{-5}$ in lymphoblastoid cell lines correspond to an estimated genome-wide false discovery rate 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 these genes, 28 were from the previously implicated set (**Supplementary Table 3**), and the other 10 genes were 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 GWAS.

34. Hyndman, R.J. & Fan, Y. Sample quantiles in statistical packages. *Am. Stat.* **50**, 361–365 (1996).
35. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.* **60**, 155–166 (2001).
36. Nica, A.C. *et al.* The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet.* **7**, e1002003 (2011).