Title
Identification of type 2 diabetes loci in 433,540 East Asian individuals

Journal
Nature, 582(7811)

ISSN
0028-0836

Authors
Spracklen, Cassandra N
Horikoshi, Momoko
Kim, Young Jin
et al.

Publication Date
2020-06-11

DOI
10.1038/s41586-020-2263-3

Peer reviewed

Published in final edited form as:

Nature. 2020 June ; 582(7811): 240–245. doi:10.1038/s41586-020-2263-3.

Identification of type 2 diabetes loci in 433,540 East Asian individuals

A full list of authors and affiliations appears at the end of the article.

SUMMARY

Meta-analyses of genome-wide association studies (GWAS) have identified >240 loci associated with type 2 diabetes (T2D)[1,2], however most loci have been identified in analyses of European-ancestry individuals. To examine T2D risk in East Asian individuals, we meta-analyzed GWAS data in 77,418 cases and 356,122 controls. In the main analysis, we identified 301 distinct association signals at 183 loci, and across T2D association models with and without consideration of body mass index and sex, we identified 61 loci newly implicated in T2D predisposition. Common variants associated with T2D in both East Asian and European populations exhibited strongly correlated effect sizes. New associations include signals in/near *GDAP1*, *PTF1A*, *SIX3, ALDH2,* a microRNA cluster, and genes that affect muscle and adipose differentiation[3]. At another locus, eQTLs at two overlapping T2D signals affect two genes, *NKX6-3* and *ANK1*, in

[*]**CORRESPONDANCE AND REQUESTS**: Correspondence and requests for materials should be addressed to Drs. Takashi Kadowaki (kadowaki-3im@h.u-tokyo.ac.jp), Robin G Walters (robin.walters@ndph.ox.ac.uk), Bong-Jo Kim (kbj6181@hanmail.net), Karen Mohlke (mohlke@med.unc.edu) and/or Xueling Sim (ephsx@nus.edu.sg).

AUTHOR CONTRIBUTIONS

Project coordination (including supervision of all experiments and analyses): K.L.M, and X.S. Manuscript writing: C.N.S., E.S.T, M.B., M.H., Y.J.K., K.L., K.L.M, X.S. Core and follow-up analyses including NKX6-1 functional work: C.N.S., M.H., Y.J.K., K.L., A.K.I., H.J.P., S.M.B., K.L.M, X.S. eQTL lookups for ANK1/NKX6-3: M.vdB., A.L.G. Core DIAMANTE group: J.E.B. (Hispanic), M.B. (European), D.W.B. (African-American), J.C.C. (South-Asian), A.M. (European), M.I.M. (European), M.C.Y.N. (African-American), L.E.P. (Hispanic), W.Zhang. (South-Asian), A.P.M. (Trans-ethnic), K.L.M., and X.S. Study-level data analyses:Y.T. (AASC), M.H. (BBJ), K.S. (BBJ), X.S. (BES), F.T. (CAGE-Amagasaki), F.T. (CAGE-GWAS), M.N. (CAGE-KING), C.N.S. (CHNS), K.L. (CKB), F.B. (CKB), C.N.S. (CLHNS), X.S. (DC/SP2), Y.J.K. (KBA), S.Moon. (KBA), C.H.T.T. (HKDR), Y.J.K. (KARE), J.Y. (MESA), X.G. (MESA), Y.T. (NAGAHAMA), J.Long. (SBC/SWHS), J.F.C. (SCES), X.S. (SCHS), V.J.Y.L. (SiMES), S.H.K. (SNUH), H.S.C. (SMC), C.H.C. (TWT2D) Individual cohort study design and/or principal investigators for the individual cohorts: M.Igase. (AASC), M.A. (BBJ), T.Kadowaki. (BBJ), Y.X.W. (BES), N.K. (CAGE-Amagasaki), N.K. (CAGE-GWAS), M.Y. (CAGE-KING), K.L.M. (CHNS), R.G.W. (CKB), K.L.M. (CLHNS), E.S.T. (DC/SP2), B.J.K. (KBA), R.C.W.M. (HKDR), B.J.K. (KARE), J.I.R. (MESA), F.M. (NAGAHAMA), X.O.S. (SBC/SWHS), C.Y.C. (SCES), W.P.K. (SCHS), T.Y.W. (SiMES), K.S.P. (SNUH), Y.S.C. (SMC), W.H.H.S. (TaiChi-G), J.Y.W. (TWT2D) Genotyping and/or phenotyping of the studies: K.K. (AASC), A.T. (BBJ), Y.K. (BBJ), T.Y. (BBJ), Y.O. (BBJ), J.B.J. (BES), T.Katsuya. (CAGE-Amagasaki), M.Isono. (CAGE-GWAS), S.I. (CAGE-KING), K.Yamamoto. (CAGE-KING), A.H. (CHNS), S.D. (CHNS), W.H. (CHNS), J.S. (CHNS), P.G.L. (CHNS), C.Y. (CKB), Y.G. (CKB), Z.B. (CKB), J.Lv. (CKB), L.L. (CKB), Z.C. (CKB), N.R.L. (CLHNS), L.S.A. (CLHNS), J.Liu. (DC/SP2), R.M.vD. (DC/SP2), S.H. (KBA), K.Yoon. (KBA), H.M.J. (KBA), D.M.S. (KBA), G.J. (HKDR), A.O.L. (HKDR), B.T. (HKDR), W.Y.S. (HKDR), J.C.N.C. (HKDR), M.Y.H. (KARE), Y.D.I.C. (MESA), T.Kawaguchi. (NAGAHAMA), Y.B.X. (SBC/SWHS), W.Zheng. (SBC/SWHS), L.Z. (SCES), C.C.K. (SCES), M.A.P. (SCHS), M.G. (SCHS), J.M.Y. (SCHS), C.S. (SiMES), M.L.C. (SiMES), M.S.L. (SMC), C.M.H. (TaiChi-G), L.M.C. (TaiChi-G), Y.J.H. (TaiChi-G), L.C.C. (TWT2D), Y.T.C. (TWT2D), F.J.T. (TWT2D), J.I.R. (MESA).

[#]Current affiliations: Stanford University, CA, USA

[^]Current affiliations: Genentech, San Francisco, CA, USA

COMPETING INTERESTS

The authors declare no competing interest.

DATA AVAILABILITY

Summary-level statistics are publicly available on the AGEN consortium website (https://blog.nus.edu.sg/agen/summary-statistics/t2d-2020), and the Accelerating Medicines Partnership T2D portal (http://www.kp4cd.org/dataset_downloads/t2d). A complete list of web resources is available in the Supplementary Information.

different tissues[4–6]. Association studies in diverse populations identify additional loci and elucidate disease genes, biology, and pathways.

Type 2 diabetes (T2D) is a common metabolic disease primarily caused by insufficient insulin production and/or secretion by the pancreatic β cells and insulin resistance in peripheral tissues[7]. Most genetic loci associated with T2D have been identified in populations of European (EUR) ancestry, including a recent meta-analysis of genome-wide association studies (GWAS) of nearly 900,000 individuals of European ancestry that identified >240 loci influencing T2D risk[1]. Differences in allele frequency between ancestries affect the power to detect associations within a population, particularly among variants rare or monomorphic in one population but more frequent in another[2,8]. Although smaller than studies in European populations, a recent T2D meta-analysis in almost 200,000 Japanese individuals identified 28 additional loci[2]. The relative contributions of different pathways to T2D pathophysiology may also differ between ancestry groups. For example, in East Asian (EAS) populations, T2D prevalence is greater than in European populations among people of similar body mass index (BMI) or waist circumference[9]. To identify new genetic associations and provide insight into T2D pathogenesis, we performed the largest meta-analysis of East Asian individuals to date.

RESULTS

We conducted a fixed-effect inverse-variance weighted GWAS meta-analysis combining 23 studies imputed to the 1000 Genomes Phase 3 reference panel from the Asian Genetic Epidemiology Network (AGEN) consortium (Supplementary Tables 1–3). We performed sex-combined T2D association without BMI adjustment in 77,418 T2D cases and 356,122 controls (effective sample size, N_{eff}=211,793). For the subset of 54,481 T2D cases and 224,231 controls (N_{eff}= 135,780) with BMI available, additional analyses were performed with and without BMI adjustment in sex-combined and sex-stratified models (Extended Data Figure 1). We defined “lead” variants as the strongest T2D-associated variants with $P<5x10^{-8}$ and defined the region +/− 500 kb from the lead variant as a locus. A locus was considered novel if the lead variant was located at least 500 kb from previously reported T2D-associated variants in any ancestry.

Using summary association statistics for ~11.8 million variants, without adjustment for BMI (Extended Data Figure 1; Supplementary Tables 1–3), we identified lead variants associated with T2D at 183 loci, of which 51 were novel (Extended Data Table 1; Extended Data Figure 2; Supplementary Table 4). Lead variants at all novel loci were common (MAF≥5%; Extended Data Figure 3), except for low-frequency variants near *GDAP1* (MAF=2.4%), which regulates mitochondrial proteins and metabolic flux in skeletal muscle[10], and *PTF1A* (MAF=4.7%), which encodes a transcription factor required for pancreatic acinar cell development[11]. Lead variants met a stricter *P*-value threshold for significance based on Bonferroni correction for 11.8 million tests ($P<4.2x10^{-9}$) at 146 of the 183 loci, including 29 of the 51 novel loci.

Using GCTA[12], we identified 301 distinct association signals that met a locus-wide significance threshold of $P<1x10^{-5}$ (Supplementary Table 5), 228 of which were genome-

wide significant ($P<5x10^{-8}$). Overall, we observed 2-4 signals at 46 loci and ≥5 signals at 12 loci. Among the ten loci with the most significant meta-analysis *P*-values of association, seven contained ≥5 distinct signals (17 signals at *INS/IGF2/KCNQ1*; 7 signals at *CDKN2A/B* and *GRM8/PAX4/LEP*; 5 signals at *CDKAL1*, *HHEX/IDE, CDC123/CAMK1D*, and *TCF7L2*; Extended Data Figure 4; Supplementary Table 5). The seven signals at the *GRM8/PAX4/LEP* locus span 1.4 Mb, and no evidence of T2D association at this locus has yet been reported in non-East Asian ancestry groups[1,13] (Extended Data Figure 4C). Joint analyses confirmed independent associations (LD r^2=0.0025) at two previously reported *PAX4* missense variants[14], rs2233580 [Arg192His: risk allele frequency (RAF)=8.6%, OR=1.31, 95% CI 1.28 – 1.34, $P_{GCTA}=3.4x10^{-93}$] and rs3824004 (Arg192Ser: RAF=3.4%, OR=1.24, 95% CI 1.19-1.28, $P_{GCTA}=1.1x10^{-30}$). The association signals at this locus also include variants near *LEP*, which encodes leptin, a hormone that regulates appetite[15]; increased leptin levels are associated with obesity and T2D[16].

At the previously reported *ANK1/NKX6-3* locus[1,17], we observed three distinct T2D association signals, two of which overlap and consist of variants spanning only ~25 kb (Figure 1). Given conflicting interpretation of candidate genes[1,5,18], we compared the T2D-association signals identified in East Asian individuals to eQTLs reported at this locus in islets[1,18–20], subcutaneous adipose[6], and skeletal muscle[5]. At the strongest signal, the lead T2D-associated variant rs33981001 is in high LD with the lead *cis*-eQTL variant for *NKX6-3* in pancreatic islets (rs12549902; EAS LD r^2=0.79, EUR r^2=0.83)[18], and the T2D risk allele is associated with decreased expression of *NKX6-3* (β=−0.36, $P=6.1x10^{-7}$; Figure 1)[4]. *NKX6-3*, or NK6 homeobox 3, encodes a pancreatic islet transcription factor required for the development of alpha and β cells in the pancreas[21] and has been shown to influence insulin secretion[16]. At the second T2D-association signal, rs62508166 is in high LD with the lead *cis*-eQTL variant for *ANK1* in subcutaneous adipose tissue[19] and skeletal muscle[15] (rs516946; EAS LD r^2=0.96, EUR r^2=0.80), and the T2D risk allele is associated with increased expression of *ANK1* (subcutaneous adipose: β=0.20, $P=1.8x10^{-7}$; skeletal muscle: β=1.01, $P=2.8x10^{-22}$). *ANK1* belongs to the ankyrin family of integral membrane proteins and has been shown to affect glucose uptake in skeletal muscle, and changes in its expression level may lead to insulin resistance[22]. Together, these GWAS and eQTL signals suggest that variants within this ~25 kb region act to increase or decrease expression levels of two different genes in different tissues to increase T2D risk.

In T2D association analyses adjusted for BMI, we identified an additional six loci, four of which were not reported previously for T2D, including loci near *MYOM3/SRSF10*, *TSN*, *GRB10*, and *NID2* (Supplementary Table 4). At the *NID2* locus, the T2D risk allele is very rare or monomorphic in non-East Asian individuals and has previously demonstrated significant associations with lower BMI and higher triglycerides in East Asian individuals, consistent with a lipodystrophy phenotype[23,24]. The lead *GRB10* variant is in low LD (EUR r^2=0.08, EAS r^2=0.57) with a variant associated with glucose-stimulated insulin secretion in European individuals[25].

Across the models with and without adjustment for BMI, correlation for the effect sizes genome-wide was higher in East Asian individuals (r=0.98) than in European individuals (r=0.89). For the 189 T2D-associated loci in East Asian individuals, the correlation

increased to 0.99 (Extended Data Figure 5). Loci with larger effects in BMI-adjusted models include *FGFR2* and *NID2*, identified only in East Asian populations and associated with lipodystrophy traits or body fat distribution. These results may reflect the role of body fat distribution in insulin resistance and T2D among East Asian individuals.

In sex-stratified analyses of males (28,027 cases and 89,312 controls) and females (27,370 cases and 135,055 controls), we identified six additional novel sex-specific loci: (i) three male-specific loci near *FOXK1*, *PDE3A*, and *IFT81*, and one female-specific locus near *LMTK2* in models without adjustment for BMI, and (ii) one male-specific locus near *LINC00851* and one female-specific locus near *CPS1* in models with adjustment for BMI (Supplementary Table 6). The lead *CPS1* variant rs1047891 (Thr1412Asn) has been reported to have a stronger effect in females than in males for cardiovascular disease and several blood metabolites[26]. Taken together, we identified a total of 61 novel loci across BMI-unadjusted, BMI-adjusted, and sex-stratified models, of which 33 met a stricter *P*-value threshold ($P<4.2x10^{-9}$).

Among all T2D-associated loci, a region spanning ~2 Mb near *ALDH2* exhibited the strongest differences between sexes (rs12231737, $P_{het}=2.6x10^{-19}$), with compelling evidence of association in males ($P_{males}=5.8x10^{-27}$) and no evidence for association in females ($P_{females}=0.19$) (Extended Data Figure 6; Supplementary Table 6). This sex difference is also observed after adjusting for BMI ($P_{males\_adjBMI}=5.2x10^{-21}$, $P_{females\_adjBMI}=0.053$). Further, joint conditional analyses revealed two conditionally distinct signals (rs12231737, $P_{GCTA}=1.7x10^{-21}$; rs557597782, $P_{GCTA}=4.9x10^{-7}$) in males only. *ALDH2* encodes aldehyde dehydrogenase 2 family member, a key enzyme in alcohol metabolism that converts acetaldehyde into acetic acid. This stretch of T2D associations in males reflects a long LD block that arose due to a recent selective sweep in East Asian individuals and results in flushing, nausea, and headache following alcohol consumption[27]. The most significantly associated missense variant in moderate LD with rs12231737 ($r^2=0.68$) was rs671 (*ALDH2* Glu504Lys: RAF=77.7%, OR=1.17, 95% CI 1.14 – 1.20, $P_{males}=1.5x10^{-24}$), which leads to reduced ALDH2 activity and reduced alcohol metabolism, and has been associated with cardiometabolic traits in East Asian populations. The T2D risk allele is associated with better tolerance for alcohol; increased BMI, blood pressure, and high-density lipoprotein cholesterol; and decreased low-density lipoprotein cholesterol and cardiovascular risk[28–30]. The strong sexual dimorphism observed at this locus may be due to differences in alcohol consumption patterns between males and females[28,30], effects of alcohol on BMI, and/or differences in the effect of alcohol on insulin sensitivity[31].

With an effective sample size comparable to the largest study of T2D in European individuals (East Asian $N_{eff}=211,793$; European $N_{eff}=231,436$)[1] and imputation to a dense 1000 Genomes reference panel, our results provide the most comprehensive and precise catalogue of East Asian T2D effects to date for comparisons across ancestries (Figure 2; Supplementary Table 7). For 183 EAS T2D loci and 231 EUR T2D loci (unadjusted for BMI)[1], we compared the per-allele effect sizes for the 332 variants available in both datasets (i.e. polymorphic and passed quality control), including lead variants from both ancestries at shared signals. Overall, the per-allele effect sizes between the two ancestries were

moderately correlated (r=0.55; Figure 2A). When the comparison was restricted to the 278 variants that are common (MAF≥5%) in both ancestries, the effect size correlation increased to r=0.59 (Figure 2B; Extended Data Figure 7). This effect size correlation further increased to r=0.87 for 106 variants significantly associated with T2D (P<5x10^{-8}) in both ancestries. Based on Cochran's heterogeneity test, 28 of 332 variants (8.4%) exhibited significant heterogeneity in effect sizes between East Asian and European populations, including 22 that were significant in only one population (Supplementary Table 7) and six with larger effect sizes in one population (e.g. *CDKAL1*, *KCNQ1*, and *HNF1B*). While the overall effect sizes for all 332 variants appear, on average, to be stronger in East Asian individuals than European individuals, this trend is reduced when each locus is represented only by the lead variant from one population (Extended Data Figure 8). Specifically, 39 variants identified in the European meta-analysis with imputation using the Haplotype Reference Consortium panel are missing from the comparison because they were rare/monomorphic or poorly imputed in the East Asian meta-analysis, with imputation based on the smaller and more heterogenous 1000 Genomes reference panel.

Variants exhibiting the largest differences in effect sizes across ancestries are generally rare (MAF ≤0.1%) in European populations but common (e.g. *PAX4*, *RANBP3L*) or low-frequency (e.g. *ZNF257*, *DGKD*) in East Asian populations. For example, rs142395395 near *ZNF257* (RAF=96.9%, OR=1.24, 95% CI 1.19-1.29, P=7.0x10^{-23}) has been reported only twice in 15,414 individuals of non-Finnish European ancestry from the gnomAD database[32]. This variant tags a previously described inversion of 415 kb observed only in East Asian individuals that disrupts the coding sequence and expression of *ZNF257*, as well as lymphoblastoid expression of 81 downstream genes and transcripts[33]. These data suggest that *ZNF257* and/or downstream target genes influence T2D susceptibility.

We identified many loci for which the lead variants exhibited similar allele frequencies and effect sizes in both the East Asian and European meta-analyses, but only reached genome-wide significance in the East Asian meta-analysis. Given shared susceptibility across ancestry groups, these loci may be detected in non-East Asian populations when sample sizes increase. Among these variants is rs117624659, located near *NKX6-1* (P_{EAS} = 2.0x10^{-16}, P_{EUR} =2.2x10^{-4}). This lead variant overlaps a highly conserved region that shows open chromatin specific to pancreatic islets. We conducted transcriptional reporter assays in MIN6 mouse insulinoma cells and observed that rs117624659 exhibited significant allelic differences in enhancer activity (Figure 3). In the pancreas, NK6 homeobox 1 (NKX6.1) is required for the development of insulin-producing β cells and is a potent bifunctional transcriptional regulator[34]. Further, inactivation of *Nkx6.1* in mice demonstrated rapid-onset diabetes due to defects in insulin biosynthesis and secretion[35]. Unexpectedly, the T2D risk allele showed increased transcriptional activity, suggesting that the variant does not act in isolation or that *NXK6-1* is not the target gene.

At one of the novel T2D-associated loci near *SIX3*, the risk allele of East Asian lead variant rs12712928-C (RAF=40.2%, OR=1.06, 95% CI 1.04 – 1.07, P=1.8x10^{-14}) is common across non-East Asian ancestries, ranging from 16.0% in Europeans to 26.4% in South Asians; however, there was no evidence of association in the other ancestry groups (meta-analysis: OR=0.98, 95% CI 0.96 – 0.99, P=2.9x10^{-3}; Extended Data Figure 9A,

Supplementary Table 8). Within the East Asian meta-analysis, the direction of effect is consistent across East Asian countries (Extended Data Figure 9B) and within the contributing cohorts (Extended Data Figure 9C). The T2D risk allele rs12712928-C is associated with higher fasting glucose levels in East Asian populations, has the strongest association with lower expression levels of both *SIX3* and *SIX2* in pancreatic islets[19], and demonstrated allele-specific binding to the transcription factor GABPA and significantly lower levels of transcriptional activity[36]. While rs12712928-C is present on only one common haplotype in most populations, it is present on an additional common haplotype (frequency =0.075) in East Asians, suggesting that the effect size attributed to rs12712928 may be influenced by other nearby unknown variants.

To identify potential candidate genes underlying the T2D-association signals identified in East Asian individuals, we further characterized 92 known and novel loci for which the lead variant at the primary East Asian association signal is located >500 kb from the lead variant of any European T2D association signal[2] (Supplementary Table 9). We characterized loci using prior trait associations, *cis*-regulatory effects on expression (colocalized eQTL), predicted effects on protein sequence, and a literature search (Supplementary Tables 10–13). Based on association results from cardiometabolic trait consortia[37], Biobank Japan[38], and the UK Biobank[39], the lead T2D-associated variant at 18 of the 88 loci was associated ($P<5x10^{-8}$) with at least one additional cardiometabolic trait, most frequently BMI or a fat mass trait (15 loci; Supplementary Tables 10 and 12). At 12 of the examined loci, T2D signals were colocalized with *cis*-eQTLs for transcripts in subcutaneous adipose tissue (n=5), skeletal muscle (n=3), pancreas (n=2), islets (n=3), or blood (n=5; Supplementary Tables 11–12), generating hypotheses of target genes and directions of effect; further examination of these candidate genes is warranted. At 19 loci, the lead T2D-associated variant or a proxy (East Asian $r^2>0.80$) alter the protein sequence (Supplementary Tables 12). These variants affect mesenchymal stem cell differentiation and adipogenesis (*GIT2*, *STEAP2* and *JMJD1C*), muscle stem cell biology (*CALCR*), glucose metabolism (*PGM1* and *SCTR*), and insulin secretion (*FGFR4*; Supplementary Table 13). At *SCTR*, which encodes the G-protein coupled secretin receptor, the lead variant encodes Ala122Pro, located in the hormone receptor domain. While mechanistic inference is required, these potential molecular mechanisms suggest new T2D susceptibility genes primarily detected by analyses in East Asian individuals.

T2D loci were also identified at clusters of noncoding RNAs with roles in islet β cell function. One locus includes a set of microRNAs specifically expressed in islet β cells, the maternally expressed noncoding RNA *MEG3,* and the paternally expressed gene *DLK1.* Targets of these microRNAs increase β cell apoptosis[40], and reduced *Meg3* impairs insulin secretion[41]. *DLK1* inhibits adipocyte differentiation, protecting from obesity[3], and promotes pancreatic ductal cell differentiation into β cells, increasing insulin secretion[42,43]. Other variants near *MEG3* have been associated with type 1 diabetes (EAS and EUR LD $r^2=0$ with EAS lead variant)[44]. The other noncoding RNA locus is the *MIR17HG* cluster of miRNAs that regulate glucose-stimulated insulin secretion and pancreatic β cell proliferation stress[45]; one of these microRNAs, miR-19a, affects hepatic gluconeogenesis[46]. Yet another T2D locus is located near *TRAF3*, which is a direct target of the *MIR17HG* microRNA cluster

and promotes hyperglycemia by increasing hepatic glucose production[47,48]. The T2D association results suggest that these noncoding RNAs influence disease susceptibility.

DISCUSSION

These T2D GWAS meta-analyses in the largest number of East Asian individuals analyzed to date identified 61 novel loci, providing additional insight into the biological basis of T2D. The results emphasize substantial shared T2D susceptibility with European individuals, as shown by the strong correlation of effect sizes among T2D-associated genetic variants with common allele frequencies in both East Asian and European ancestry populations. Compared to a recent T2D study in individuals of European ancestry[1], we observed less attenuation of effects on T2D in analyses adjusted for BMI. Loci with a greater effect on T2D after adjusting for BMI include loci with lipodystrophy-like traits identified only in East Asian individuals to date, adding support to the observation[49,50] that factors beyond overall BMI, such as visceral adiposity or lipodystrophy, may also play a role in T2D in East Asians. The results also detect novel associations in East Asian individuals identified because they have higher allele frequencies in East Asian populations, exhibit larger effect sizes, and/or are influenced by other environmental or behavioral factors such as alcohol consumption.

The identified loci point to multiple plausible molecular mechanisms and many new candidate genes linking T2D susceptibility to diverse biological processes. Following the annotation of loci identified in the East Asian meta-analysis, we speculate a substantial role for insulin resistance in T2D pathogenesis among East Asian individuals through skeletal muscle, adipose, and liver development and function. We also provide evidence that multiple distinct association signals in the same region do not necessarily act through the same gene. Conditionally distinct association signals in close proximity can affect different genes that may act in different tissues by different mechanisms, emphasizing the value of identifying functional variants that enable variant-to-gene links to be examined directly. Our results provide a foundation for future biological research in T2D pathogenesis and offer potential targets for mechanisms for interventions in disease risk.

METHODS

Ethics statement

All human research was approved by the relevant institutional review boards for each study at their respective sites (Supplementary Table 1) and conducted according to the Declaration of Helsinki. All participants provided written informed consent.

Study cohorts and quality control

The East Asian type 2 diabetes (T2D) meta-analyses were performed with studies participating in the Asian Genetic Epidemiology Network (AGEN), a consortium of genetic epidemiology studies of T2D and related traits conducted in individuals of East Asian ancestry, and the Diabetes Meta-analysis of Trans-ethnic Association Studies (DIAMANTE), a consortium examining the genetic contribution to T2D across diverse ancestry populations including African-American, East Asian, European, Hispanic, and

South Asian. The East Asian meta-analysis included 77,418 T2D cases and 356,122 controls from 23 GWAS, including three biobanks, CKB, KBA[51,52], and BBJ[2] [effective sample size (N_{eff}) = 211,793; Extended Data Figure 1]. A subset of studies with BMI measurement available was analyzed with and without BMI adjustment in sex-combined and sex-specific models (54,481 cases, 224,231 controls; N_{eff} = 135,780). For each study, T2D case control ascertainment is described in Supplementary Table 1 and summary statistics are provided in Supplementary Table 2. As T2D case definitions across cohorts differ, it is possible that cases of type 1 diabetes and maturity onset diabetes of the young (MODY) are included in these meta-analyses. Included studies were genotyped on either commercially available or customized Affymetrix or Illumina genome-wide genotyping arrays. Array quality control criteria implemented within each study, including variant call rate and Hardy-Weinberg equilibrium, are summarized in Supplementary Table 3. To harmonize study-level genotype scaffold for imputation to 1000 Genomes (1000G) reference panels, each study adopted a uniform protocol for pre-imputation quality checks. Each study applied the protocol to exclude variants with: i) mismatched chromosomal positions or alleles not present in the reference panel; ii) ambiguous alleles (AT/CG) with minor allele frequency (MAF) >40% in the reference panel; or iii) absolute allele frequency differences >20% compared to East Asian-specific allele frequencies. The genotype scaffold for each study was then imputed to the 1000G Phase 1 or 3 reference panel[53] using minimac3[54] or IMPUTEv2[55]. In BMI-unadjusted analyses, all studies were imputed to 1000G Phase 3. In BMI-adjusted and sex-stratified analyses, all studies were imputed to 1000G Phase 3 except for a subset of Biobank Japan[17], which was imputed to the 1000G Phase 1 reference panel.

Study-level association analyses

Within each study, all variants were tested for association with T2D assuming an additive model of inheritance within a regression framework, including age, sex, and other study-specific covariates (Supplementary Table 3). To account for population structure and relatedness, association analyses were either performed using FIRTH[56] or mach2dat with additional adjustment for principal components in unrelated individuals or a linear mixed model with kinship matrix implemented in BOLT-LMM[57]. In studies analyzed with the linear mixed model, allelic effects and standard errors were converted to the log-odds scale that accounts for case-control imbalance[58]. Within each study, variants were removed if i) imputation quality score was poor (minimac3 r^2<0.30; IMPUTE2 info score <0.40); ii) combined case control minor allele count <5; or iii) standard error of the log-OR>10. For a subset of the studies, BMI was added as an additional covariate, and association analyses were also performed separately in males and females. For each study and model, association statistics were corrected with genomic control inflation factor[59] calculated from common variants (MAF≥5%) (Supplementary Table 3). For BBJ, we applied the genomic control inflation factor 1.21 as reported[2].

Sex-combined meta-analysis

We combined study-level association statistics using fixed-effects meta-analysis with inverse-variance weighting of log-ORs implemented in METAL[60]. Variants with allele frequency differences >20% between 1000G Phase 1 and 3 panels were excluded from the meta-analysis. To assess excess inflation arising from cryptic relatedness and population

structure, we applied LD score regression to the meta-analysis summary statistics to estimate residual inflation of summary statistics, using a set of 1,889 unrelated Chinese individuals from the Singapore Chinese Eye Study[61]. The LD score regression intercepts were 0.993 for BMI-unadjusted, and 1.0163 for BMI-adjusted models. As the LD score regression intercepts indicated absence of excess inflation, the meta-analysis results were corrected for inflation using these LD score regression intercepts. For subsequent analyses, we considered only variants that were present in at least 50% of the effective sample size N_{eff} [computed as $4/(1/N_{cases} + 1/N_{controls})$][60]. Heterogeneity in allelic effect sizes between studies were assessed with fixed-effects inverse-variance weighted meta-analysis P_{het}. We further compared the genetic effects from BMI-unadjusted and BMI-adjusted models using fixed-effects inverse-variance weighted meta-analysis P_{het}. Loci were defined as novel if the lead variant is: (1) at least 500 kb away and confirmed by GCTA to be conditionally independent from previously reported T2D-associated variants in any ancestry, and (2) assessed using LocusZoom plots and detailed literature review to be away from known loci with extended LD. Lead variants mapping to loci already associated with other glycemic traits were still considered novel for the association with T2D.

BMI adjustment analyses and effect size comparison

For the subset of studies with both BMI-unadjusted and BMI-adjusted models, we compared the effect sizes and heterogeneity of the lead variants using the standardized mean difference to account for the correlation between the two models[1]. We calculated the Pearson correlation coefficient between effect sizes from the BMI-unadjusted and BMI-adjusted models for all 13.2M variants genome-wide (r=0.98) and for the lead variants at 189 T2D-associated loci (r=0.99).

Sex-differentiated meta-analysis

The meta-analyses described above were repeated for males and females separately. The male-specific meta-analyses included up to 28,027 cases and 89,312 controls (N_{eff} = 65,660) and the female-specific analyses included up to 27,370 cases and 135,055 controls (N_{eff} = 70,332). LD score regression intercepts were 1.0044 for BMI-unadjusted and 1.0045 for BMI-adjusted models in males and 1.0050 for BMI-unadjusted and 1.0187 for BMI-adjusted models in females. We further performed a test for heterogeneity in allelic effects between males and females as implemented in GWAMA[62,63].

Detection of distinct association signals

To detect multiple distinct association signals at each associated locus, we combined overlapping loci when the distance between any pair of lead variants was <1 Mb. We then performed approximate conditional analyses using GCTA[12] with genome-wide meta-analysis summary statistics and LD estimated from 78,000 samples from the Korean Biobank Array[52]. We note the limitations in using a single population reference panel for LD estimation for a meta-analysis of diverse East Asian populations. We present all distinct signals at conditional threshold of $P<1x10^{-5}$, but we suggest that readers exhibit caution and limit inferences from these analyses to signals that show the strongest evidence of association.

Comparing loci effects between East Asian and European populations

We compared the per-allele effect sizes of lead variants identified from the East Asian BMI-unadjusted sex-combined meta-analysis (183 lead variants) and European BMI-unadjusted sex-combined meta-analysis[1] (231 lead variants; Supplementary Table 7). Across the 414 associated variants from the two ancestries, 12 lead variants overlapped, resulting in 402 unique variants. As the variants in the European analysis were imputed using the Haplotype Reference Consortium reference panel and did not include indel variants, a variant in strong LD (East Asian r^2>0.90) with the lead East Asian variant was used when the lead variant was an indel, when possible. If the lead East Asian variant or a variant in strong LD (East Asian r^2>0.90) was not available in the European data from DIAMANTE, we used results from a previous European type 2 diabetes meta-analysis[64]. The effect size comparison plot was restricted to 332 variants where data was available for both ancestries (Figure 2A). For loci that were significant in both the East Asian and European meta-analyses, if the lead variants were different, both lead variants were plotted (see Supplementary Table 7). Effect size plots were further restricted to: i) 278 lead variants with MAF≥5% in both East Asian and European meta-analyses (Extended Data Figure 7); ii) 203 lead variants significant in the East Asian meta-analysis (Extended Data Figure 8A); and iii) 234 lead variants significant in the European meta-analysis (Extended Data Figure 8B). Differences in effect sizes between the two populations could be due to differences in imputation quality with different reference panels.

Associations with other metabolic traits and outcomes

We examined publicly available GWAS summary statistics (mostly available through the Type 2 Diabetes Knowledge Portal[37]) to explore associations of the lead variant at the 92 loci for which there are no genome-wide significant European variants within 500 kb (listed in Supplementary Table 9). Association statistics from the following consortia were available for query on the portal (last accessed August 28, 2019): coronary artery disease from CARDIoGRAM[65], BMI and waist-hip-ratio from GIANT[66,67], lipid traits from GLGC[68], and glycemic traits from MAGIC[69–73]. Additionally, we used available data from AGEN East Asian meta-analyses for lipids[74], along with the phenotypic data from the UK Biobank[75], BioBank Japan[24,38], and blood pressure data from ICBP[76]. For this analysis, we looked up the effect size and *P*-value of the East Asian lead variants in the other datasets. If the variant-trait association reached at least nominal significance (P<1x10^{-3}), we included the lookup results in Supplementary Table 10. When the lead East Asian variant was missing in the prior GWAS data, we reported it as “NF” (not found) in the table.

Colocalization with expression quantitative trait loci (eQTL)

We searched publicly available eQTL databases such as GTEx v7[77] and the Parker lab Islet Browser[19], to identify *cis*-eQTLs at the novel loci in adipose (subcutaneous and visceral), blood, pancreas, pancreatic islet, and skeletal muscle tissue. We also searched for *cis*-eQTLs in subcutaneous adipose tissue data from the METSIM study[6], whole blood[78], and peripheral blood (BioBank Japan; http://jenger.riken.jp/en/result). Colocalized eQTLs were identified if the lead expression level-associated variant and the GWAS lead variant were in high LD (r^2>0.80) in Europeans to accommodate the predominantly European eQTL data.

Reciprocal conditional analyses were also performed using the METSIM data to determine if the GWAS lead variant and the lead expression-associated variant were part of the same eQTL signal.

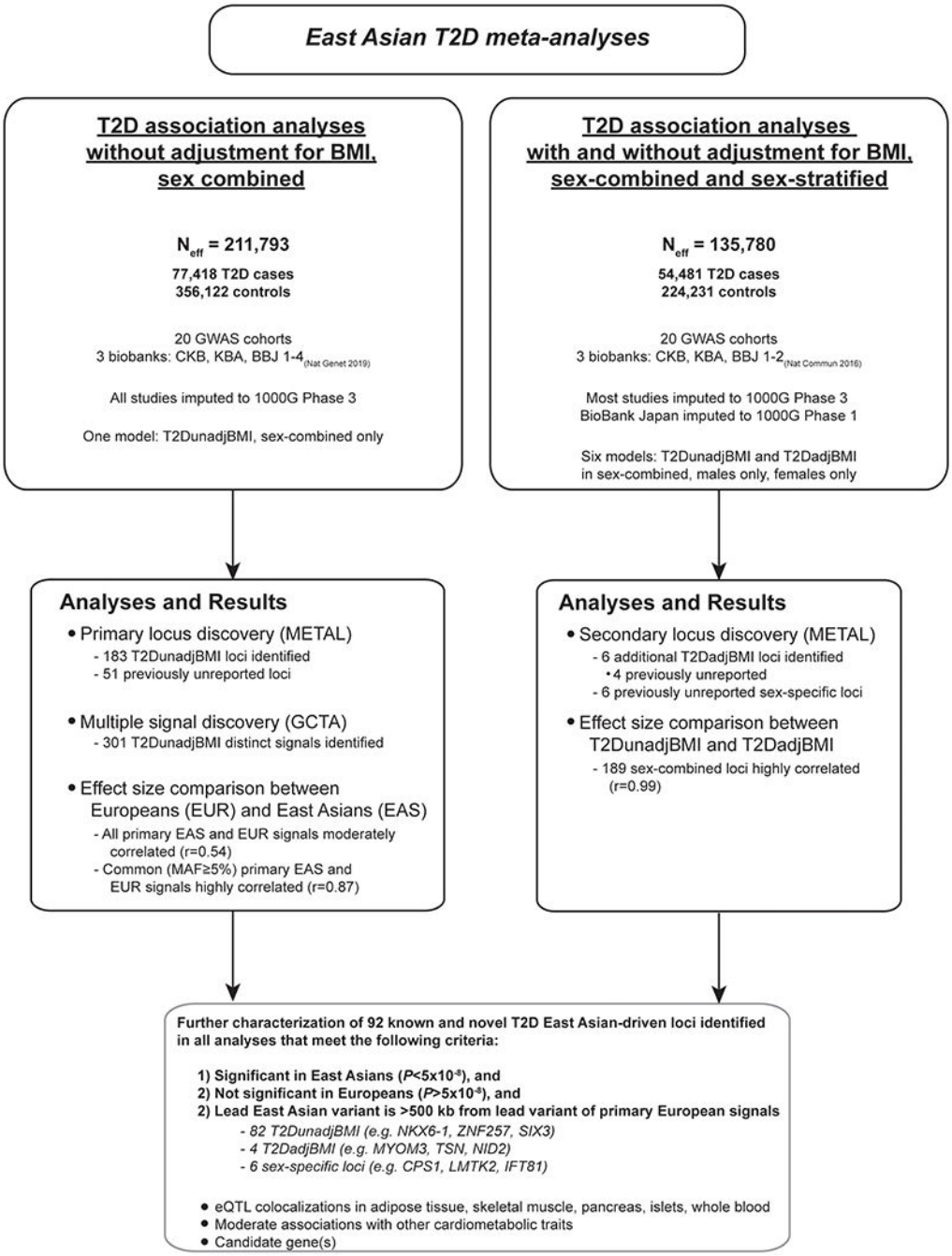
Literature review

We conducted a traditional literature review to identify candidate genes at each novel locus using NCBI Entrez Gene, PubMed and OMIM. We included gene symbols and the following keywords as search terms in PubMed: diabetes, glucose, insulin, islet, adipose, muscle, liver, obesity. A gene was considered a potential candidate if an apparent link to T2D biology existed based on prior studies of gene function.

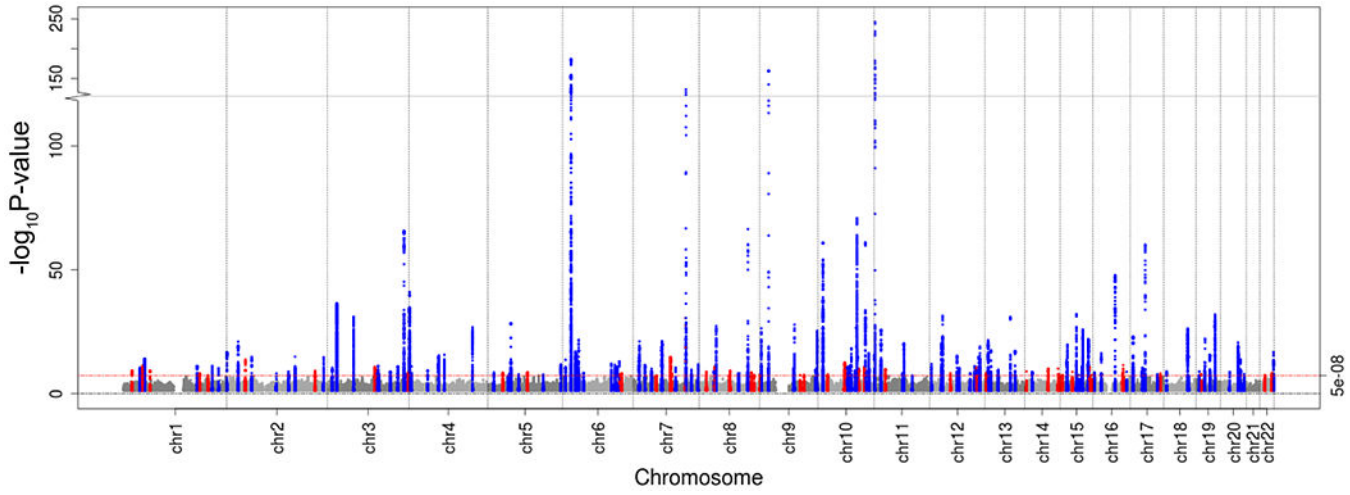
Functional annotation and experimentation at *NKX6-1*

We used ENCODE[79], ChromHMM[80], and Human Epigenome Atlas[81] data available through the UCSC Genome Browser to identify candidate variants at the association signal near *NKX6-1* that overlapped open-chromatin peaks, ChromHMM chromatin states, and chromatin-immunoprecipitation sequencing (ChIP-seq) peaks of histone modifications H4K4me1, H3K4me3, and H3K27ac, and transcription factors in the pancreas and pancreatic islets. MIN6 mouse insulinoma cells (FROM ATCC)[82] were cultured in DMEM (Sigma) supplemented with 10% FBS, 1mM sodium pyruvate, and 0.1 mM beta-mercaptoethanol. The cell cultures were maintained at 37° C with 5% CO_2. To measure variant allelic differences in enhancer activity at the *NKX6*-1 locus, we designed oligonucleotide primers (forward: CCCTAGTAATGCCCTTTTTGTT; reverse: TCAGCCTGAGAAGTCTGTGA) with KpnI and Xhol restriction sites, and amplified the 400-bp DNA region (GRCh37/hg19 -chr4: 85,339,430-85,339,829) around rs117624659. As previously described[80], we ligated amplified DNA from individuals homozygous for each allele into the multiple cloning site of the pGL4.23 (Promega) minimal promoter luciferase reporter vector in both the forward and reverse orientations with respect to the genome. Clones were isolated and sequenced for genotype and fidelity. $2.1x10^5$ MIN6 cells were seeded per well and grown to 90% confluence in 24-well plates. We co-transfected five independent luciferase constructs and *Renilla* control reporter vector (phRL-TK, Promega) using Lipofectamine 2000 (Life Technologies) and incubated. 48-hours post-transfection, the cells were lysed with Passive Lysis Buffer (Promega). Luciferase activity was measured using the Dual-luciferase Reporter Assay System (Promega) per manufacturer instructions and as previously described[83]. MIN6 cell lines were authenticated through genotyping and tested negative for mycoplasma contamination.

Extended Data

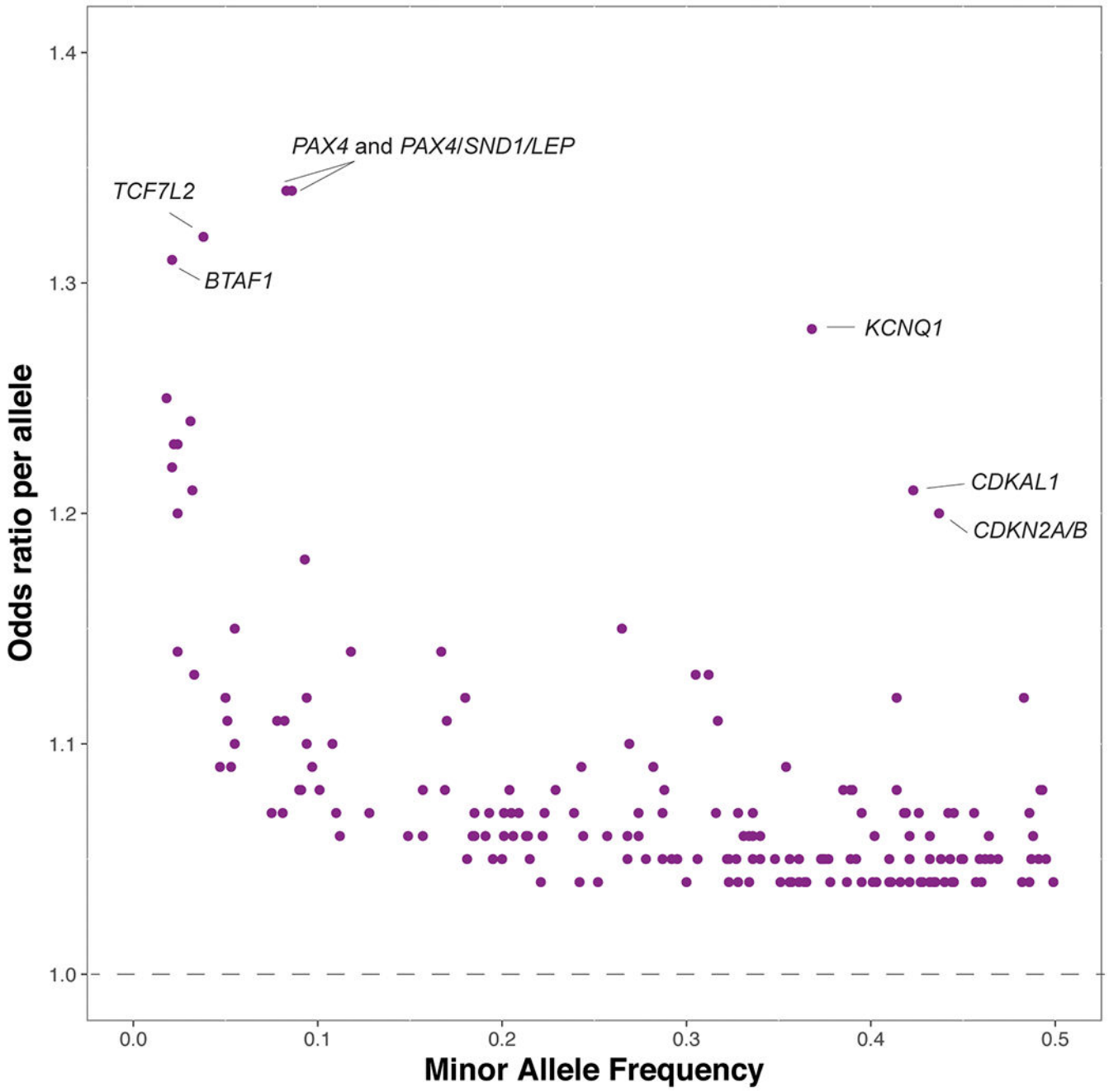


Extended Data Figure 1:
Flow chart of study design, depicting the different data analyses performed.



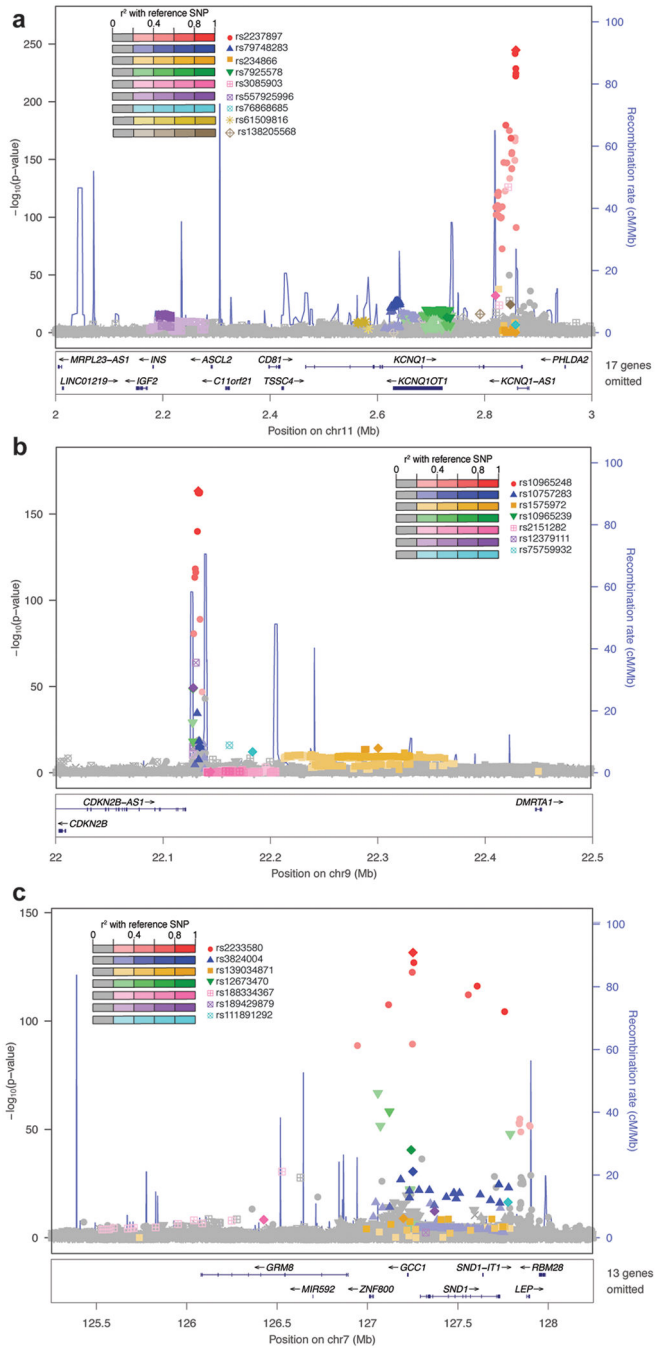
Extended Data Figure 2: Manhattan plot for East Asian T2D meta-analysis association results in model unadjusted for BMI.

$-\log_{10}(P)$ values from two-sided fixed-effects inverse-variance genome-wide meta-analysis association results for each variant (*y*-axis; maximal N_{eff}=211,793) was plotted against the genomic position (hg19; *x*-axis). Known T2D loci achieving genome-wide significance ($P<5.0\text{x}10^{-8}$) meta-analysis are shown in blue. Loci achieving genome-wide significance that are previously unreported for T2D association are shown in red.



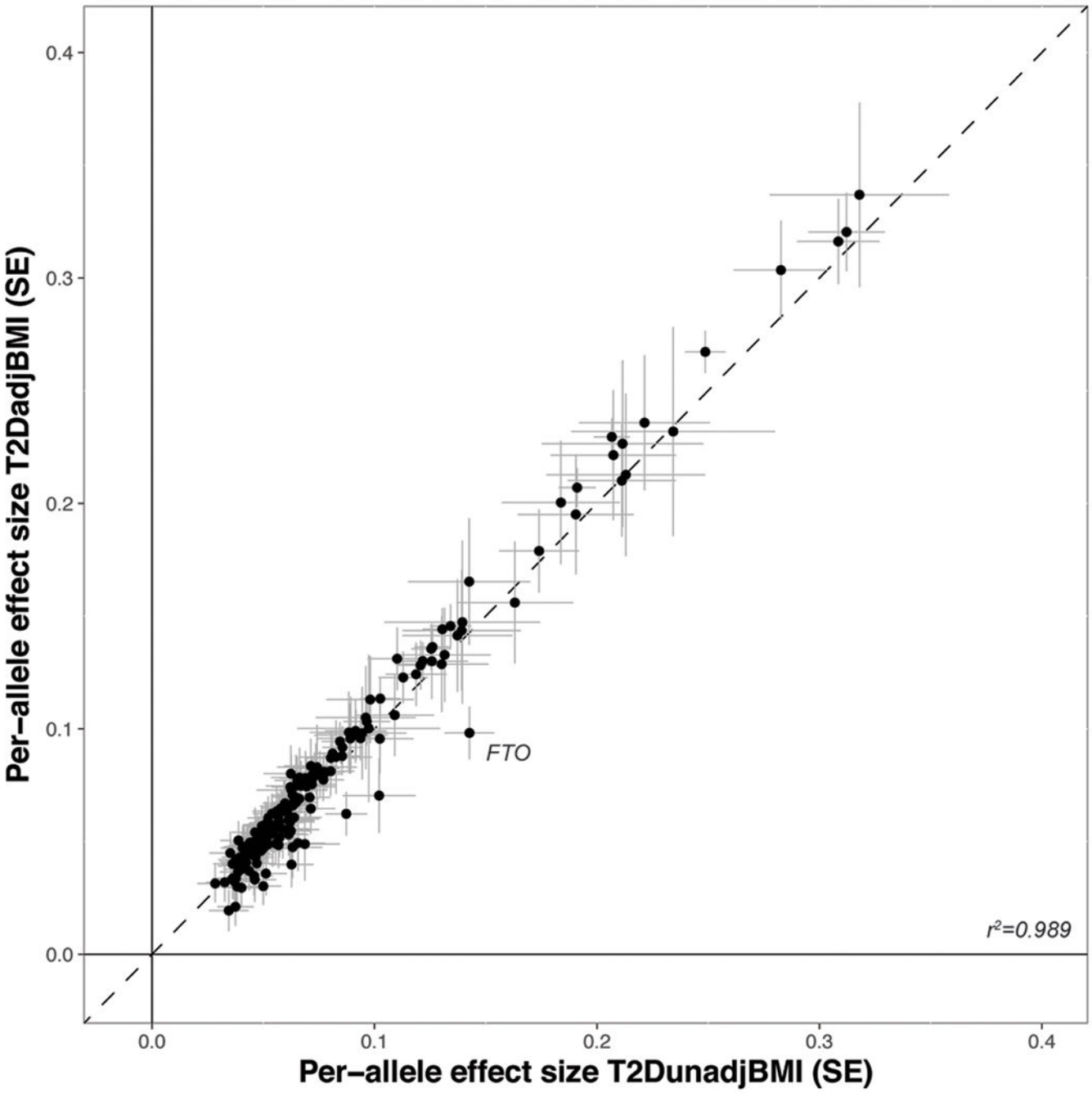
Extended Data Figure 3: The relationship between effect size and minor allele frequency.

Odds ratios (*y*-axis) and minor allele frequencies (*x*-axis) for 189 primary association signals from the T2D BMI-unadjusted models. Odds ratios are from two-sided fixed-effects inverse-variance meta-analysis on a maximal effective sample size of 211,793.



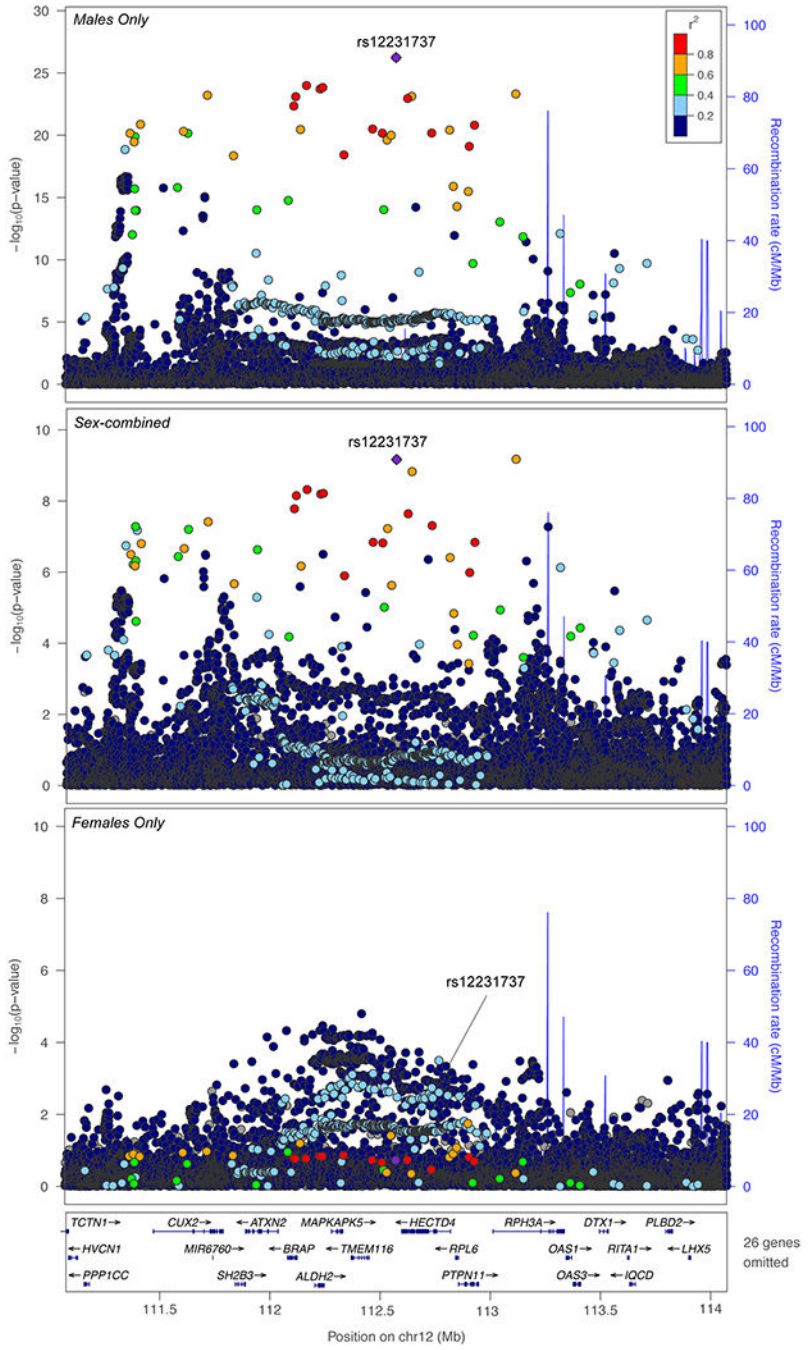
Extended Data Figure 4: Regional association plots at three T2D-associated loci with the strongest association *P*-value and more than five distinct association signals in East Asians.

(A) *INS/IGF2/KCNQ1*, **(B)** *CDKN2A/B*, **(C)** *PAX4/LEP*. $-\log_{10}(P)$ values were from the two-sided fixed-effect inverse-variance meta-analysis. Distinct signals ($P<1.0\times10^{-6}$ from GCTA conditional analyses) were plotted; N_{eff} for each distinct signal are reported in Supplementary Table 4. Variants are colored based on East Asian 1000G Phase 3 LD with the lead variants for each association signal, shown as diamonds.



Extended Data Figure 5: Effect size comparison of lead variants in sex-combined models unadjusted and adjusted for BMI.

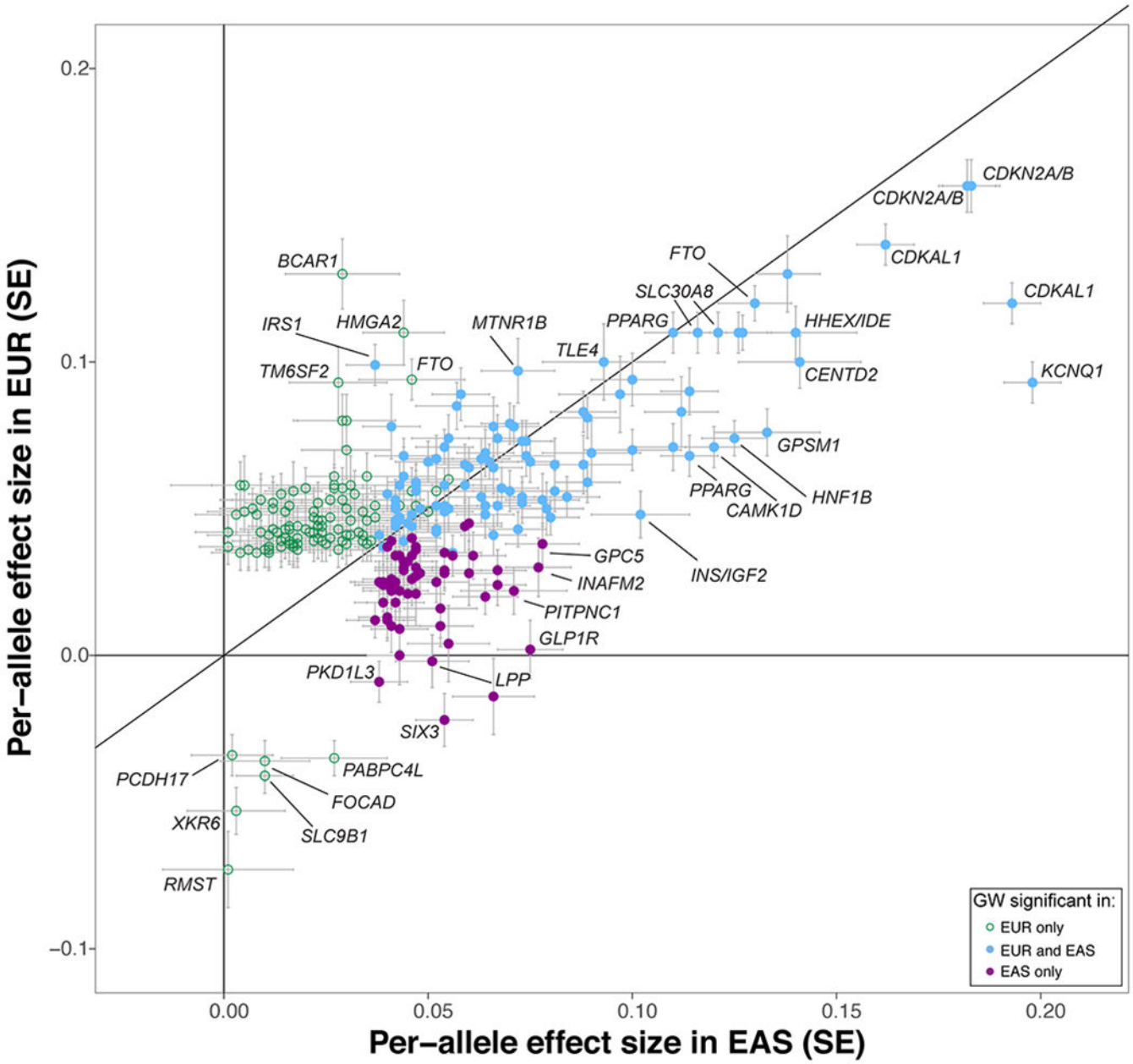
At 189 lead variants identified in the East Asian BMI-unadjusted sex-combined T2D meta-analysis, per-allele effect sizes (β) from the BMI-adjusted sex-combined model were plotted against the BMI-unadjusted sex-combined model. Both sex-combined models were from two-sided fixed-effect inverse-variance meta-analyses and included the same set of studies for comparable sample size. Each point denotes the per-allele effect size; standard errors of the effect size estimates extend out as grey lines. Effect sizes between the two models are highly correlated with a Pearson correlation coefficient r=0.99 (Supplementary Table 4).



Extended Data Figure 6: Regional plots of male-specific T2D-associated locus, *ALDH2*.

For each plot, $-\log_{10}(P)$ values from association results from two-sided fixed-effect inverse-variance meta-analyses for each variant (*y*-axis) was plotted against the genomic position (hg19; *x*-axis). The lead variant rs12231737 plotted is the lead variant from the BMI-unadjusted male-specific meta-analysis (N_{eff}=65,202) and also the sex-combined meta-analysis (N_{eff}=138,947) from the same subset of individuals included in the sex-stratified analyses (female-specific N_{eff}=70,051). This lead variant rs12231737 is in high LD with rs77768175, identified from the larger BMI-unadjusted sex-combined meta-analysis (East

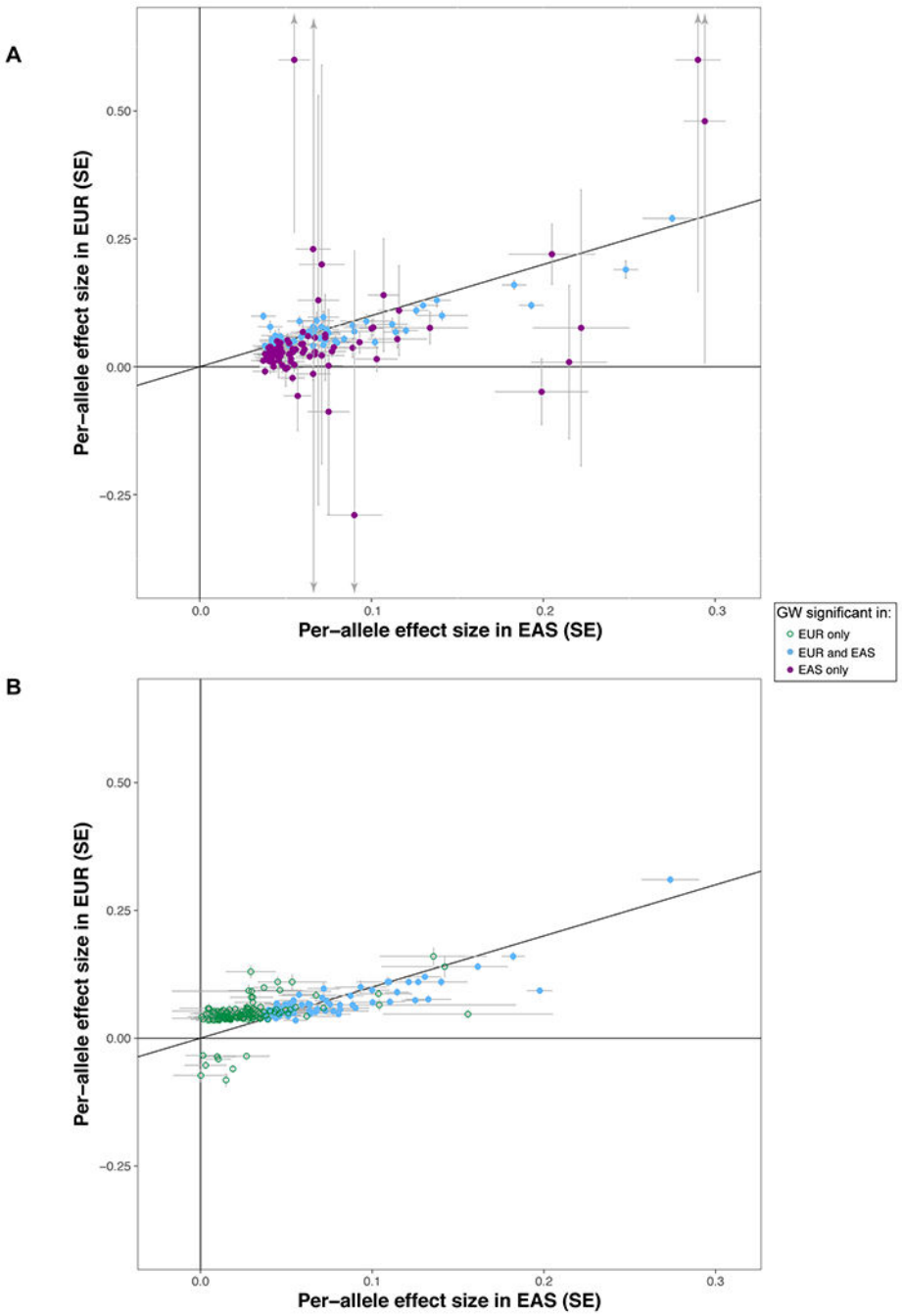
Asian r^2=0.80). **(A)** Males only, **(B)** sex-combined, and **(C)** females only. Variants are shaded based on East Asian 1000G Phase 3 LD with the lead variant, shown as a purple diamond.



Extended Data Figure 7: Effect size comparison of common lead variants (MAF≥5%) identified in this East Asian meta-analysis and a previously published European T2D GWAS meta-analysis[2].

For 278 unique lead variants with MAF≥5% in both the East Asian and European BMI-unadjusted meta-analyses, per-allele effect sizes (β) from Mahajan et al.[2] (*y*-axis) were plotted against per-allele effect sizes from this East Asian meta-analysis (*x*-axis). Effect sizes from both meta-analyses were from two-sided fixed-effect inverse-variance meta-analyses (maximal N_{eff}=211,793 for East Asian and 231,436 for European meta-analyses). Each point denotes the per-allele effect size; standard errors of the effect size estimates extend out as grey lines. Variants are colored purple if they were significant in the East Asian meta-analysis only, green if they were significant in European meta-analysis only, and

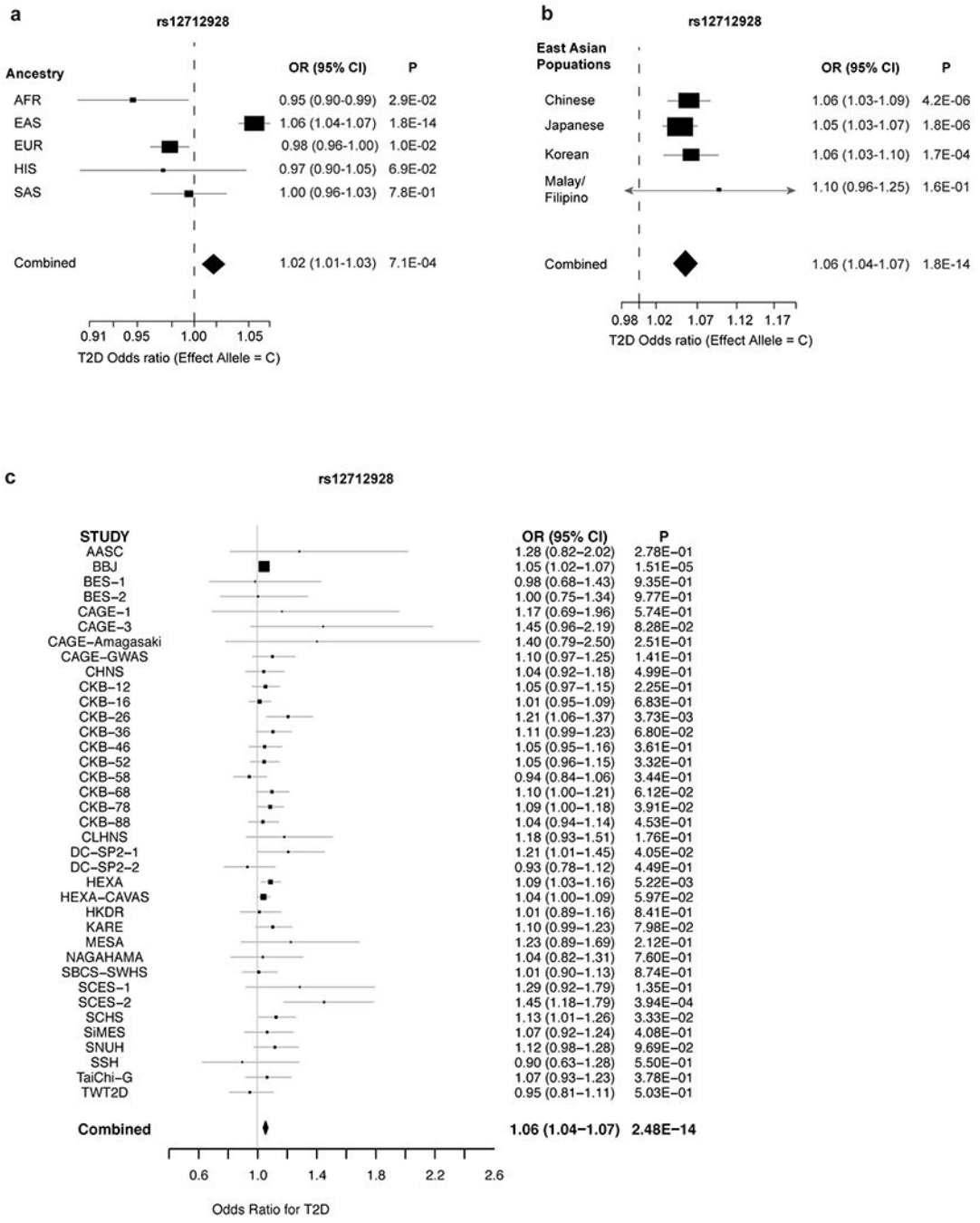
blue if they were significant in both the East Asian and European meta-analyses. (see Methods and Supplementary Table 7).



Extended Data Figure 8: Effect size comparison of lead variants identified in East Asian BMI-unadjusted meta-analysis and previously published European T2D GWAS meta-analysis[2].

For 332 lead variants identified from the two BMI-unadjusted meta-analyses, per-allele effect sizes (β) from a European meta-analysis (*y*-axis) were plotted against per-allele effect sizes from this East Asian meta-analysis (*x*-axis). Effect sizes from both meta-analyses were from two-sided fixed-effect inverse-variance meta-analysis (maximal N_{eff}=211,793 for East Asian and 231,436 for European meta-analyses). Each point denotes the per-allele effect size; standard errors of the effect size estimates extend out as grey lines. **(A)** 152 lead

variants significant in the East Asian meta-analysis (purple) or both the East Asian and European meta-analysis (blue) and **(B)** 192 lead variants significant in the European meta-analysis (green) or both the East Asian and European meta-analysis (blue). These plots include only one variant per locus, in contrast to Figure 2 and Extended Data Figure 7.



Extended Data Figure 9: Forest plots of BMI-unadjusted meta-analysis association results at *SIX3-SIX2* locus.

Odds ratios (black boxes) and 95% confidence intervals (horizontal lines extending out) for T2D associations at the lead East Asian variant (rs12712928) are presented **(A)** across ancestries of African-American (AFR), East Asian (EAS), European (EUR)[2], Hispanic (HIS), and South Asian (SAS) individuals, **(B)** within four major East Asian populations (Chinese, Japanese, Korean, and Malay/Filipino combined due to small sample sizes), **(C)** from each contributing cohort. Effect sizes from East Asian study, ancestry, population, and

combined meta-analysis were from two-sided fixed-effect inverse-variance meta-analysis. The size of the box is proportional to the sample size of each contributing study/ancestry/population, which are available in Supplementary Table 8. This East Asian study had >90% power to detect the observed association with a MAF=0.40, OR=1.06, and 77,418 T2D cases. Given the number of T2D cases and frequency of rs12712928-C within the other datasets, at 80% power, we can reasonably exclude association OR >1.07 in EUR and >1.15 in AFR, HIS, and SAS between rs12782928 and T2D. Full study names can be found in Supplementary Table 1 and corresponding sample sizes can be found in Supplementary Table 2.

Extended Data Table 1:

Novel lead variants associated with type 2 diabetes in East Asians.

Single variant association results from East Asian fixed-effect inverse-variance meta-analysis (BMI-unadjusted sex-combined model adjusted for age, sex, and study-specific covariates) using METAL. Loci were defined as novel if the lead variant is (1) at least 500 kb away and confirmed by GCTA to be conditionally independent from previously reported T2D-associated variants in any ancestry, and (2) assessed using locuszoom plots and biology lookups to be to be away from known loci with extended LD. Four additional variants met the definition for a novel locus but are located within the previously reported major histocompatibility complex (MHC) region; see Supplementary Table S5 for the full list of distinct association signals at the MHC region. rs4804181 was >500kb from primary signal rs3111316 in European meta-analysis, but <500kb from their secondary signal rs755734872. Genome-wide significant association is defined as P < 5.0E-08. Physical position based on hg19. Effect alleles are associated with increased risk for T2D. Odds ratios reflect per allele effects of variants on T2D risk.

Chr, chromosome; Pos, position; RAF, risk allele frequency; Neff, effective sample size; OR, odds ratio; CI, confidence interval; P, P-value

Locus	Lead variant	Chr	Pos	RA	NRA	RAF	Cases (N)	Controls (N)	Neff	OR (95% CI)	*P*	P_{het}
VWA5B1	rs60573766	1	20,688,352	C	T	0.64	77,418	356,122	211,793	1.04 (1.03-1.06)	4.3E-10	8.2E-01
MAST2	rs562138031	1	46,244,900	C	CT	0.73	77,221	351,786	211,039	1.06 (1.04-1.07)	4.0E-12	7.5E-01
PGM1	rs2269245	1	64,107,893	G	A	0.82	77,221	351,786	211,039	1.06 (1.04-1.07)	5.4E-10	3.9E-01
TSEN15	rs1327123	1	184,014,593	C	G	0.46	77,418	356,122	211,793	1.04 (1.03-1.05)	7.0E-09	9.2E-01
MDM4	rs201297151	1	204,474,581	CAAAAAAAA	C	0.44	77,221	351,786	211,039	1.04 (1.03-1.05)	3.4E-08	3.7E-01
SIX3	rs12712928	2	45,192,080	C	G	0.40	77,418	356,122	211,793	1.06 (1.04-1.07)	1.8E-14	2.4E-01
IKZF2	rs75179644	2	213,687,103	T	C	0.90	77,418	356,122	211,793	1.08 (1.05-1.10)	5.4E-10	7.6E-01
ZBTB20	rs6806156	3	114,968,018	T	C	0.61	77,418	356,122	211,793	1.05 (1.03-1.06)	1.6E-11	8.7E-01
TFRC	rs9866168	3	195,830,310	T	A	0.64	77,418	356,122	211,793	1.05 (1.03-1.06)	1.6E-09	2.7E-01
RANBP3L	rs16902871	5	36,257,018	G	A	0.15	77,418	356,122	211,793	1.06 (1.04-1.08)	3.3E-09	2.0E-01
PCSK1	rs6556925	5	95,848,503	C	A	0.42	77,418	356,122	211,793	1.04 (1.03-1.05)	3.1E-09	2.2E-01
REPS1	rs9376382	6	139,205,386	C	T	0.60	77,418	356,122	211,793	1.04 (1.03-1.05)	1.5E-08	5.4E-01
HIVEP2	rs9390022	6	143,056,556	T	C	0.80	77,418	356,122	211,793	1.05 (1.03-1.07)	6.4E-09	7.2E-01
ZNF713	rs565050730	7	55,984,953	GA	G	0.33	77,418	356,122	211,793	1.04 (1.03-1.06)	4.4E-08	7.1E-01
STEAP1	rs62469016	7	89,752,238	C	G	0.22	77,418	356,122	211,793	1.07 (1.05-1.08)	1.5E-15	2.4E-01
CALCR	rs2074120	7	93,107,093	A	C	0.32	77,418	356,122	211,793	1.04 (1.03-1.06)	8.4E-09	9.3E-02
GRM8/PAX4	rs117737118	7	126,526,991	G	A	0.09	77,062	350,162	210,460	1.18 (1.14-1.21)	3.3E-31	3.3E-01
ASAH1	rs34642578	8	17,927,609	T	C	0.05	77,418	356,122	211,793	1.09 (1.06-1.13)	1.6E-09	5.2E-02
ZNF703	rs4739515	8	37,391,203	G	C	0.54	77,418	356,122	211,793	1.05 (1.03-1.06)	1.7E-11	7.5E-01
FGFR1	rs328301	8	38,343,012	T	C	0.33	77,418	356,122	211,793	1.04 (1.03-1.06)	4.1E-08	6.9E-01
KCNB2	rs349359	8	73,503,743	C	A	0.24	77,418	356,122	211,793	1.04 (1.03-1.06)	3.1E-08	2.5E-01
GDAP1	rs149265787	8	75,214,398	G	A	0.02	77,392	355,608	211,694	1.14 (1.10-1.19)	5.7E-10	1.1E-01
TRIB1	rs60089934	8	126,471,274	A	G	0.38	77,418	356,122	211,793	1.04 (1.03-1.06)	3.3E-09	1.0E-01
EFR3A	rs73708054	8	132,879,795	C	T	0.25	77,418	356,122	211,793	1.04 (1.03-1.06)	4.4E-08	2.3E-01
DMRT2	rs1016565	9	1,032,567	A	G	0.42	77,418	356,122	211,793	1.04 (1.02-1.05)	2.2E-08	3.4E-01
PTCH1	rs113154802	9	98,278,413	C	T	0.89	77,418	356,122	211,793	1.06 (1.04-1.08)	3.5E-08	7.5E-02
ABCA1	rs201375651	9	107,597,527	CA	C	0.39	77,418	356,122	211,793	1.04 (1.03-1.06)	2.6E-08	2.0E-01
PTF1A	rs77065181	10	23,487,778	A	G	0.05	77,418	356,122	211,793	1.09 (1.06-1.13)	1.6E-08	8.3E-01
ARID5B	rs141583966	10	63,712,602	G	GGTGT	0.91	77,175	351,384	210,874	1.08 (1.05-1.11)	7.7E-10	2.6E-01
JMJD1C	rs148928116	10	64,976,133	T	TA	0.79	77,418	356,122	211,793	1.06 (1.04-1.08)	2.5E-13	8.7E-02
ARHGAP19	rs10736116	10	99,056,921	C	G	0.31	77,418	356,122	211,793	1.05 (1.03-1.06)	9.2E-11	7.7E-01
BBIP1	rs7895872	10	112,678,657	T	G	0.58	77,418	356,122	211,793	1.05 (1.03-1.06)	1.4E-11	5.5E-01
BDNF	rs4922793	11	27,729,505	A	G	0.57	77,418	356,122	211,793	1.04 (1.03-1.06)	1.6E-10	3.8E-01
FAIM2	rs77978149	12	50,269,863	T	C	0.09	77,259	354,498	211,214	1.08 (1.05-1.10)	5.7E-09	2.1E-01
ALDH2	rs149212747	12	111,836,771	A	AC	0.79	77,221	351,786	211,039	1.07 (1.05-1.09)	2.1E-11	1.7E-05
RBM19	rs7307263	12	114,123,722	G	C	0.43	77,418	356,122	211,793	1.04 (1.02-1.05)	3.6E-08	9.6E-01
FGF9	rs9316706	13	22,589,883	A	G	0.35	77,418	356,122	211,793	1.04 (1.03-1.06)	3.3E-09	6.5E-01
NYNRIN	rs12437434	14	24,878,370	C	T	0.71	77,221	351,786	211,039	1.05 (1.03-1.06)	1.0E-09	3.8E-01
LRRC74A	rs58524310	14	77,382,503	G	A	0.33	75,090	353,783	207,126	1.05 (1.03-1.06)	8.4E-11	3.7E-01
DLK1/MEG3/ miRNA cluster	rs73347525	14	101,255,172	A	G	0.76	74,931	352,159	206,547	1.06 (1.04-1.08)	7.5E-11	3.8E-01
TRAF3	rs55700915	14	103,237,952	A	G	0.43	77,221	351,786	211,039	1.04 (1.03-1.06)	1.5E-08	9.9E-01
HERC2	rs76704029	15	28,546,173	T	C	0.73	66,648	265,839	174,183	1.06 (1.04-1.08)	3.4E-08	1.5E-01
MYO5C	rs149336329	15	52,587,740	G	T	0.95	77,221	351,786	211,039	1.11 (1.07-1.14)	1.7E-09	8.3E-01
RGMA	rs61021634	15	93,825,384	A	G	0.44	77,418	356,122	211,793	1.05 (1.03-1.06)	1.4E-11	1.1E-01
IGF1R	rs79826452	15	99,366,409	A	G	0.89	77,418	356,122	211,793	1.07 (1.05-1.10)	3.2E-08	4.1E-01
PKD1L3	rs12600132	16	72,022,534	T	C	0.43	77,418	356,122	211,793	1.04 (1.03-1.05)	6.0E-09	7.9E-01
ZFHX3	rs6416749	16	73,100,308	C	T	0.37	77,062	350,162	210,460	1.05 (1.04-1.07)	3.4E-12	5.9E-01
SUMO2	rs35559984	17	73,187,031	CA	C	0.65	74,931	352,159	206,547	1.05 (1.03-1.07)	7.9E-09	9.8E-01
ZNF799	rs4604181	19	12,509,536	A	C	0.51	77,062	350,162	210,460	1.04 (1.03-1.06)	1.5E-08	5.5E-01
ZNRF3	rs147413364	22	29,380,119	T	TA	0.36	77,175	351,384	210,874	1.04 (1.03-1.06)	3.4E-08	1.4E-01
WNT7B	rs28637892	22	46,313,618	T	G	0.22	63,772	319,376	175,881	1.05 (1.04-1.07)	3.7E-09	4.2E-01

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Cassandra N Spracklen[1,2], Momoko Horikoshi[3], Young Jin Kim[4], Kuang Lin[5], Fiona Bragg[5], Sanghoon Moon[4], Ken Suzuki[3,6,7,8], Claudia HT Tam[9,10], Yasuharu Tabara[11], Soo-Heon Kwak[12], Fumihiko Takeuchi[13], Jirong Long[14], Victor JY Lim[15], Jin-Fang Chai[15], Chien-Hsiun Chen[16], Masahiro Nakatochi[17], Jie Yao[18], Hyeok Sun Choi[19], Apoorva K Iyengar[1], Hannah J Perrin[1], Sarah M Brotman[1], Martijn van de Bunt[20,21], Anna L Gloyn[20,21,22,#], Jennifer E Below[23,24], Michael Boehnke[25], Donald W Bowden[26,27], John C Chambers[28,29,30,31,32], Anubha Mahajan[20,21], Mark I McCarthy[20,21,22,^], Maggie CY Ng[23,26], Lauren E Petty[23,24], Weihua Zhang[29,30], Andrew P Morris[21,33,34], Linda S Adair[35], Masato Akiyama[6,36,37], Zheng Bian[38], Juliana CN Chan[9,10,39,40], Li-Ching Chang[16], Miao-Li Chee[41], Yii-Der Ida Chen[18], Yuan-Tsong Chen[16], Zhengming Chen[5], Lee-Ming Chuang[42,43], Shufa Du[35], Penny Gordon-Larsen[35], Myron Gross[44], Xiuqing Guo[18], Yu Guo[38], Sohee Han[4], Annie-Green Howard[45], Wei Huang[46], Yi-Jen Hung[47,48], Mi Yeong Hwang[4], Chii-Min Hwu[49,50], Sahoko Ichihara[51], Masato Isono[13], Hye-Mi Jang[4], Guozhi Jiang[9,10], Jost B Jonas[52], Yoichiro Kamatani[6,53], Tomohiro Katsuya[54,55], Takahisa Kawaguchi[11], Chiea-Chuen Khor[56,57], Katsuhiko Kohara[58], Myung-Shik Lee[59,60], Nanette R Lee[61], Liming Li[62], Jianjun Liu[56,63], Andrea O Luk[9,10], Jun Lv[62], Yukinori Okada[8,64], Mark A Pereira[65], Charumathi Sabanayagam[41,66,67], Jinxiu Shi[46], Dong Mun Shin[4], Wing Yee So[9,39], Atsushi Takahashi[6,68], Brian Tomlinson[9,69], Fuu-Jen Tsai[70], Rob M van Dam[15,63], Yong-Bing Xiang[71], Ken Yamamoto[72], Toshimasa Yamauchi[7], Kyungheon Yoon[4], Canqing Yu[62], Jian-Min Yuan[73,74], Liang Zhang[41], Wei Zheng[14], Michiya Igase[75], Yoon Shin Cho[19], Jerome I Rotter[76], Ya-Xing Wang[77], Wayne HH Sheu[48,50,78], Mitsuhiro Yokota[79], Jer-Yuarn Wu[16], Ching-Yu Cheng[41,66,67], Tien-Yin Wong[41,66,67], Xiao-Ou Shu[14], Norihiro Kato[13], Kyong-Soo Park[12,80,81], E-Shyong Tai[15,63,82], Fumihiko Matsuda[11], Woon-Puay Koh[15,83], Ronald CW Ma[9,10,39,40], Shiro Maeda[3,84,85], Iona Y Millwood[5,86], Juyoung Lee[4], Takashi Kadowaki[7,*], Robin G Walters[5,86,*], Bong-Jo Kim[4,*], Karen L Mohlke[1,*], Xueling Sim[15,*]

Affiliations

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

[2]Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA USA

[3]Laboratory for Endocrinology, Metabolism and Kidney Diseases, RIKEN Centre for Integrative Medical Sciences, Yokohama, Japan

[4]Division of Genome Research, Center for Genome Science, National Institute of Health, Cheongju-si, Korea

[5]Nuffield Department of Population Health, University of Oxford, Oxford, UK

[6]Laboratory for Statistical and Translational Genetics, RIKEN Centre for Integrative Medical Sciences, Yokohama, Japan

[7]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[8]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan

[9]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China

[10]Chinese University of Hong Kong-Shanghai Jiao Tong University Joint Research Centre in Diabetes Genomics and Precision Medicine, The Chinese University of Hong Kong, Hong Kong, China

[11]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

[12]Department of Internal Medicine, Seoul National University Hospital, Seoul, South Korea

[13]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan

[14]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

[15]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore

[16]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

[17]Department of Nursing, Nagoya University Graduate School of Medicine, Nagoya University Hospital, Nagoya, Japan

[18]Department of Pediatrics, The Institute for Translational Genomics and Population Sciences, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA

[19]Biomedical Science, Hallym University, Chuncheon, South Korea

[20]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

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

[22]Oxford NIHR Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, Churchill Hospital, Oxford, UK

[23]Vanderbilt Genetics Institute, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

[24]Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA

[25]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA

[26]Center for Genomics and Personalized Medicine Research, Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC, USA

[27]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA

[28]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[29]Department of Epidemiology and Biostatistics, Imperial College London, London, UK

[30]Department of Cardiology, Ealing Hospital, London North West Healthcare NHS Trust, Middlesex, UK

[31]Imperial College Healthcare NHS Trust, Imperial College London, London, UK

[32]MRC-PHE Centre for Environment and Health, Imperial College London, London, UK

[33]Department of Biostatistics, University of Liverpool, Liverpool, UK

[34]School of Biological Sciences, University of Manchester, Manchester, UK

[35]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[36]Laboratory for Statistical Analysis, RIKEN Centre for Integrative Medical Sciences, Yokohama, Japan

[37]Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

[38]Chinese Academy of Medical Sciences, Beijing, China

[39]Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Hong Kong, China

[40]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China

[41]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore

[42]Division of Endocrinology & Metabolism, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[43]Institute of Preventive Medicine, School of Public Health, National Taiwan University, Taipei, Taiwan

[44]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA

[45]Department of Biostatistics, Carolina Population Center, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[46]Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China

[47]Division of Endocrine and Metabolism, Tri-Service General Hospital Songshan Branch, Taipei, Taiwan

[48]School of Medicine, National Defense Medical Center, Taipei, Taiwan

[49]Section of Endocrinology and Metabolism, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan

[50]School of Medicine, National Yang-Ming University, Taipei, Taiwan

[51]Department of Environmental and Preventive Medicine, Jichi Medical University School of Medicine, Shimotsuke, Japan

[52]Department of Ophthalmology, Medical Faculty Mannheim of the University of Heidelberg, Mannheim, Germany

[53]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

[54]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Osaka, Japan

[55]Department of Geriatric and General Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan

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

[57]Department of Biochemistry, National University of Singapore, Singapore, Singapore

[58]Department of Regional Resource Management, Ehime University Faculty of Collaborative Regional Innovation, Ehime, Japan

[59]Severance Biomedical Science Institute and Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea

[60]Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea

[61]Department of Anthropology, Sociology and History, University of San Carlos, Cebu City, Philippines

[62]Departments of Epidemiology & Biostatistics, Peking University Health Science Centre, Peking University, Beijing, China

[63]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore

[64]Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Osaka, Japan

[65]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

[66]Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore, Singapore

[67]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore

[68]Department of Genomic Medicine, National Cerebral and Cardiovascular Center, Osaka, Japan

[69]Faculty of Medicine, Macau University of Science and Technology, Macau, China

[70]Department of Medical Genetics and Medical Research, China Medical University Hospital, Taichung, Taiwan

[71]State Key Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

[72]Department of Medical Biochemistry, Kurume University School of Medicine, Kurume, Japan

[73]Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA, USA

[74]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

[75]Department of Anti-aging Medicine, Ehime University Graduate School of Medicine, Ehime, Japan

[76]Departments of Pediatrics and Medicine, The Institute for Translational Genomics and Population Sciences, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA

[77]Beijing Institute of Ophthalmology, Ophthalmology and Visual Sciences Key Laboratory, Beijing Tongren Hospital, Capital Medical University, Beijing, China

[78]Division of Endocrinology and Metabolism, Department of Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

[79]Kurume University School of Medicine, Kurume, Japan

[80]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, South Korea

[81]Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea

[82]Duke-NUS Medical School, Singapore, Singapore

[83]Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

[84]Department of Advanced Genomic and Laboratory Medicine, Graduate School of Medicine, University of the Ryukyus, Okinawa, Japan

[85]Division of Clinical Laboratory and Blood Transfusion, University of the Ryukyus Hospital, Okinawa, Japan

[86]Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK

[87]These authors contributed equally to this work

ACKNOWLEDGEMENTS

This work was supported by subawards (to X.S and Y.S.C.) from NIDDK U01DK105554 (Jose C. Florez). The authors thank all investigators, staff members and study participants for their contributions to all participating studies, including but not limited to: CAGE – Drs Toshio Ogihara, Yukio Yamori, Akihiro Fujioka, Chikanori Makibayashi, Sekiharu Katsuya, Ken Sugimoto, Kei Kamide, and Ryuichi Morishita; SBCS – Ms. Regina Courtney, Drs. Hui Cai, Ben Zhang and Ms. Jing He; SMC – Dr Duk-Hwan Kim. A full list of funding, and individual and study acknowledgements are available in Supplementary Information.

REFERENCES

1. Mahajan A et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 50, 1505–1513, doi:10.1038/s41588-018-0241-6 (2018). [PubMed: 30297969]
2. Suzuki K et al. Identification of 28 new susceptibility loci for type 2 diabetes in the Japanese population. Nat Genet 51, 379–386, doi:10.1038/s41588-018-0332-4 (2019). [PubMed: 30718926]
3. Moon YS et al. Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. Molecular and cellular biology 22, 5585–5592, doi:10.1128/mcb.22.15.5585-5592.2002 (2002). [PubMed: 12101250]
4. van de Bunt M et al. The miRNA profile of human pancreatic islets and beta-cells and relationship to type 2 diabetes pathogenesis. PLoS One 8, e55272, doi:10.1371/journal.pone.0055272 (2013). [PubMed: 23372846]
5. Scott LJ et al. The genetic regulatory signature of type 2 diabetes in human skeletal muscle. Nat Commun 7, 11764, doi:10.1038/ncomms11764 (2016). [PubMed: 27353450]
6. Civelek M et al. Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits. Am J Hum Genet 100, 428–443, doi:10.1016/j.ajhg.2017.01.027 (2017). [PubMed: 28257690]
7. Stumvoll M, Goldstein BJ & van Haeften TW Type 2 diabetes: principles of pathogenesis and therapy. Lancet (London, England) 365, 1333–1346, doi:10.1016/s0140-6736(05)61032-x (2005).
8. Cho YS et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet 44, 67–72, doi:10.1038/ng.1019 (2011). [PubMed: 22158537]
9. Huxley R et al. Ethnic comparisons of the cross-sectional relationships between measures of body size with diabetes and hypertension. Obesity reviews : an official journal of the International Association for the Study of Obesity 9 Suppl 1, 53–61, doi:10.1111/j.1467-789X.2007.00439.x (2008). [PubMed: 18307700]
10. Lassiter DG, Sjogren RJO, Gabriel BM, Krook A & Zierath JR AMPK activation negatively regulates GDAP1, which influences metabolic processes and circadian gene expression in skeletal muscle. Mol Metab 16, 12–23, doi:10.1016/j.molmet.2018.07.004 (2018). [PubMed: 30093355]

11. Hoang CQ et al. Transcriptional Maintenance of Pancreatic Acinar Identity, Differentiation, and Homeostasis by PTF1A. Molecular and cellular biology 36, 3033–3047, doi:10.1128/MCB.00358-16 (2016). [PubMed: 27697859]
12. Yang J et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 44, 369–375, S361–363, doi:10.1038/ng.2213 (2012). [PubMed: 22426310]
13. Fuchsberger C et al. The genetic architecture of type 2 diabetes. Nature 536, 41–47, doi:10.1038/nature18642 (2016). [PubMed: 27398621]
14. Kwak SH et al. Nonsynonymous Variants in PAX4 and GLP1R Are Associated With Type 2 Diabetes in an East Asian Population. Diabetes 67, 1892–1902, doi:10.2337/db18-0361 (2018). [PubMed: 29941447]
15. Klok MD, Jakobsdottir S & Drent ML The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. Obesity reviews : an official journal of the International Association for the Study of Obesity 8, 21–34, doi:10.1111/j.1467-789X.2006.00270.x (2007). [PubMed: 17212793]
16. Rasmussen-Torvik LJ et al. Associations of body mass index and insulin resistance with leptin, adiponectin, and the leptin-to-adiponectin ratio across ethnic groups: the Multi-Ethnic Study of Atherosclerosis (MESA). Ann Epidemiol 22, 705–709, doi:10.1016/j.annepidem.2012.07.011 (2012). [PubMed: 22929534]
17. Imamura M et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. Nat Commun 7, 10531, doi:10.1038/ncomms10531 (2016). [PubMed: 26818947]
18. van de Bunt M et al. Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. PLoS Genet 11, e1005694, doi:10.1371/journal.pgen.1005694 (2015). [PubMed: 26624892]
19. Varshney A et al. Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. Proc Natl Acad Sci U S A 114, 2301–2306, doi:10.1073/pnas.1621192114 (2017). [PubMed: 28193859]
20. Thurner M et al. Integration of human pancreatic islet genomic data refines regulatory mechanisms at Type 2 Diabetes susceptibility loci. eLife 7, doi:10.7554/eLife.31977 (2018).
21. Henseleit KD et al. NKX6 transcription factor activity is required for alpha- and beta-cell development in the pancreas. Development 132, 3139–3149, doi:10.1242/dev.01875 (2005). [PubMed: 15944193]
22. Yan R et al. A novel type 2 diabetes risk allele increases the promoter activity of the muscle-specific small ankyrin 1 gene. Sci Rep 6, 25105, doi:10.1038/srep25105 (2016). [PubMed: 27121283]
23. Wen W et al. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. 23, 5492–5504, doi:10.1093/hmg/ddu248 (2014).
24. Akiyama M et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. Nat Genet 49, 1458–1467, doi:10.1038/ng.3951 (2017). [PubMed: 28892062]
25. Prokopenko I et al. A central role for GRB10 in regulation of islet function in man. PLoS Genet 10, e1004235, doi:10.1371/journal.pgen.1004235 (2014). [PubMed: 24699409]
26. Hartiala JA et al. Genome-wide association study and targeted metabolomics identifies sex-specific association of CPS1 with coronary artery disease. Nat Commun 7, 10558, doi:10.1038/ncomms10558 (2016). [PubMed: 26822151]
27. Okada Y et al. Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. Nat Commun 9, 1631, doi:10.1038/s41467-018-03274-0 (2018). [PubMed: 29691385]
28. Xu F et al. ALDH2 genetic polymorphism and the risk of type II diabetes mellitus in CAD patients. Hypertens Res 33, 49–55, doi:10.1038/hr.2009.178 (2010). [PubMed: 19876063]

29. Kato N et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet 43, 531–538, doi:10.1038/ng.834 (2011). [PubMed: 21572416]
30. Takeuchi F et al. Confirmation of ALDH2 as a Major locus of drinking behavior and of its variants regulating multiple metabolic phenotypes in a Japanese population. Circ J 75, 911–918 (2011). [PubMed: 21372407]
31. Schrieks IC, Heil AL, Hendriks HF, Mukamal KJ & Beulens JW The effect of alcohol consumption on insulin sensitivity and glycemic status: a systematic review and meta-analysis of intervention studies. Diabetes Care 38, 723–732, doi:10.2337/dc14-1556 (2015). [PubMed: 25805864]
32. Lek M et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291, doi:10.1038/nature19057 (2016). [PubMed: 27535533]
33. Puig M et al. Functional Impact and Evolution of a Novel Human Polymorphic Inversion That Disrupts a Gene and Creates a Fusion Transcript. PLoS Genet 11, e1005495, doi:10.1371/journal.pgen.1005495 (2015). [PubMed: 26427027]
34. Iype T et al. The transcriptional repressor Nkx6.1 also functions as a deoxyribonucleic acid context-dependent transcriptional activator during pancreatic beta-cell differentiation: evidence for feedback activation of the nkx6.1 gene by Nkx6.1. Molecular endocrinology (Baltimore, Md.) 18, 1363–1375, doi:10.1210/me.2004-0006 (2004).
35. Taylor BL, Liu FF & Sander M Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells. Cell Rep 4, 1262–1275, doi:10.1016/j.celrep.2013.08.010 (2013). [PubMed: 24035389]
36. Spracklen CN et al. Identification and functional analysis of glycemic trait loci in the China Health and Nutrition Survey. PLoS Genet 14, e1007275, doi:10.1371/journal.pgen.1007275 (2018). [PubMed: 29621232]
37. Type 2 Diabetes Knowledge Portal, <http://www.type2diabetesgenetics.org/home/portalHome> (2019).
38. Kanai M et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet 50, 390–400, doi:10.1038/s41588-018-0047-6 (2018). [PubMed: 29403010]
39. Bycroft C et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 562, 203–209, doi:10.1038/s41586-018-0579-z (2018). [PubMed: 30305743]
40. Kameswaran V et al. Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. Cell metabolism 19, 135–145, doi:10.1016/j.cmet.2013.11.016 (2014). [PubMed: 24374217]
41. You L et al. Downregulation of Long Noncoding RNA Meg3 Affects Insulin Synthesis and Secretion in Mouse Pancreatic Beta Cells. J Cell Physiol 231, 852–862, doi:10.1002/jcp.25175 (2016). [PubMed: 26313443]
42. Wang Y et al. Overexpression of Pref-1 in pancreatic islet beta-cells in mice causes hyperinsulinemia with increased islet mass and insulin secretion. Biochem Biophys Res Commun 461, 630–635, doi:10.1016/j.bbrc.2015.04.078 (2015). [PubMed: 25918019]
43. Rhee M et al. Preadipocyte factor 1 induces pancreatic ductal cell differentiation into insulin-producing cells. Sci Rep 6, 23960, doi:10.1038/srep23960 (2016). [PubMed: 27044861]
44. Onengut-Gumuscu S et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nat Genet 47, 381–386, doi:10.1038/ng.3245 (2015). [PubMed: 25751624]
45. Chen Y et al. MicroRNA-17–92 cluster regulates pancreatic beta-cell proliferation and adaptation. Mol Cell Endocrinol 437, 213–223, doi:10.1016/j.mce.2016.08.037 (2016). [PubMed: 27568466]
46. Dou L et al. MiR-19a mediates gluconeogenesis by targeting PTEN in hepatocytes. Mol Med Rep 17, 3967–3971, doi:10.3892/mmr.2017.8312 (2018). [PubMed: 29257352]
47. Chen Z et al. Hepatocyte TRAF3 promotes insulin resistance and type 2 diabetes in mice with obesity. Molecular metabolism 4, 951–960, doi:10.1016/j.molmet.2015.09.013 (2015). [PubMed: 26909311]

48. Liu F, Cheng L, Xu J, Guo F & Chen W miR-17–92 functions as an oncogene and modulates NF-kappaB signaling by targeting TRAF3 in MGC-803 human gastric cancer cells. Int J Oncol 53, 2241–2257, doi:10.3892/ijo.2018.4543 (2018). [PubMed: 30226589]
49. Ma RC & Chan JC Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States. Ann N Y Acad Sci 1281, 64–91, doi:10.1111/nyas.12098 (2013). [PubMed: 23551121]
50. Zhu Y et al. Racial/Ethnic Disparities in the Prevalence of Diabetes and Prediabetes by BMI: Patient Outcomes Research To Advance Learning (PORTAL) Multisite Cohort of Adults in the U.S. Diabetes Care, doi:10.2337/dc19-0532 (2019).
51. Kim Y, Han BG & Ko G. E. S. g. Cohort Profile: The Korean Genome and Epidemiology Study (KoGES) Consortium. Int J Epidemiol 46, e20, doi:10.1093/ije/dyv316 (2017). [PubMed: 27085081]
52. Moon S et al. The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. Sci Rep 9, 1382, doi:10.1038/s41598-018-37832-9 (2019). [PubMed: 30718733]
53. Auton A et al. A global reference for human genetic variation. Nature 526, 68–74, doi:10.1038/nature15393 (2015). [PubMed: 26432245]
54. Das S et al. Next-generation genotype imputation service and methods. Nat Genet 48, 1284–1287, doi:10.1038/ng.3656 (2016). [PubMed: 27571263]
55. Howie B, Marchini J & Stephens M Genotype imputation with thousands of genomes. G3 (Bethesda) 1, 457–470, doi:10.1534/g3.111.001198 (2011). [PubMed: 22384356]
56. Ma C, Blackwell T, Boehnke M & Scott LJ Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. Genet Epidemiol 37, 539–550, doi:10.1002/gepi.21742 (2013). [PubMed: 23788246]
57. Loh PR et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet 47, 284–290, doi:10.1038/ng.3190 (2015). [PubMed: 25642633]
58. Cook JP, Mahajan A & Morris AP Guidance for the utility of linear models in meta-analysis of genetic association studies of binary phenotypes. Eur J Hum Genet 25, 240–245, doi:10.1038/ejhg.2016.150 (2017). [PubMed: 27848946]
59. Devlin B & Roeder K Genomic control for association studies. Biometrics 55, 997–1004 (1999). [PubMed: 11315092]
60. Willer CJ, Li Y & Abecasis GR METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191, doi:10.1093/bioinformatics/btq340 (2010). [PubMed: 20616382]
61. Bulik-Sullivan BK et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47, 291–295, doi:10.1038/ng.3211 (2015). [PubMed: 25642630]
62. Magi R, Lindgren CM & Morris AP Meta-analysis of sex-specific genome-wide association studies. Genet Epidemiol 34, 846–853, doi:10.1002/gepi.20540 (2010). [PubMed: 21104887]
63. Magi R & Morris AP GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics 11, 288, doi:10.1186/1471-2105-11-288 (2010). [PubMed: 20509871]
64. Scott RA et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes 66, 2888–2902, doi:10.2337/db16-1253 (2017). [PubMed: 28566273]
65. Schunkert H et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 43, 333–338, doi:10.1038/ng.784 (2011). [PubMed: 21378990]
66. Shungin D et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature 518, 187–196, doi:10.1038/nature14132 (2015). [PubMed: 25673412]
67. Yengo L et al. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. Hum Mol Genet 27, 3641–3649, doi:10.1093/hmg/ddy271 (2018). [PubMed: 30124842]
68. Willer CJ et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 45, 1274–1283, doi:10.1038/ng.2797 (2013). [PubMed: 24097068]
69. 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, doi:10.1038/ng.520 (2010). [PubMed: 20081858]

70. Saxena R et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet 42, 142–148, doi:10.1038/ng.521 (2010). [PubMed: 20081857]
71. Strawbridge RJ et al. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. Diabetes 60, 2624–2634, doi:10.2337/db11-0415 (2011). [PubMed: 21873549]
72. Manning AK et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 44, 659–669, doi:10.1038/ng.2274 (2012). [PubMed: 22581228]
73. Wheeler E et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. PLoS medicine 14, e1002383, doi:10.1371/journal.pmed.1002383 (2017). [PubMed: 28898252]
74. Spracklen CN et al. Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. Hum Mol Genet 27, 1122, doi:10.1093/hmg/ddx439 (2018). [PubMed: 29351605]
75. Sudlow C et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine 12, e1001779, doi:10.1371/journal.pmed.1001779 (2015). [PubMed: 25826379]
76. Ehret GB et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 478, 103–109, doi:10.1038/nature10405 (2011). [PubMed: 21909115]
77. Gamazon ER et al. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. Nat Genet 50, 956–967, doi:10.1038/s41588-018-0154-4 (2018). [PubMed: 29955180]
78. Võsa U et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. bioRxiv, 447367, doi:10.1101/447367 (2018).
79. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74, doi:10.1038/nature11247 (2012). [PubMed: 22955616]
80. Ezzat S et al. The cancer-associated FGFR4-G388R polymorphism enhances pancreatic insulin secretion and modifies the risk of diabetes. Cell Metab 17, 929–940, doi:10.1016/j.cmet.2013.05.002 (2013). [PubMed: 23747250]
81. Kundaje A et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330, doi:10.1038/nature14248 (2015). [PubMed: 25693563]
82. Miyazaki J et al. Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. Endocrinology 127, 126–132, doi:10.1210/endo-127-1-126 (1990). [PubMed: 2163307]
83. Fogarty MP, Cannon ME, Vadlamudi S, Gaulton KJ & Mohlke KL Identification of a regulatory variant that binds FOXA1 and FOXA2 at the CDC123/CAMK1D type 2 diabetes GWAS locus. PLoS Genet 10, e1004633, doi:10.1371/journal.pgen.1004633 (2014). [PubMed: 25211022]

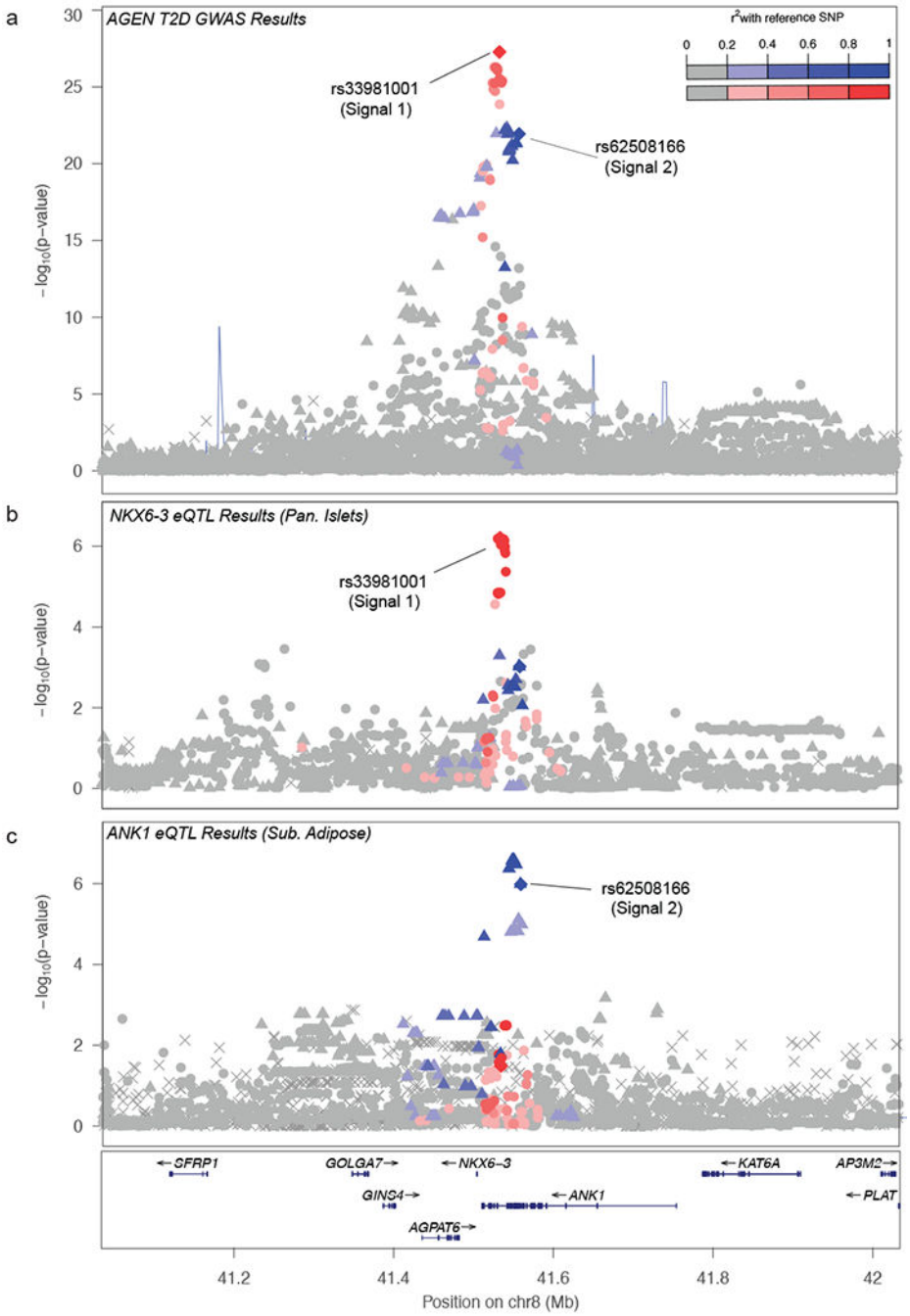


Figure 1: Two distinct T2D-association signals at the *ANK1-NKX6-3* locus associated with expression levels of two transcripts in two tissues.

(A) Regional association plot for East Asian sex-combined BMI-unadjusted two-sided fixed-effect inverse-variance meta-analysis at *ANK1-NKX6-3* locus. Approximate conditional analysis using GTCA identified three distinct T2D-association signals (P<1x10^{-5}) at this locus (signal 1, rs33981001, N_{eff}=211,793; signal 2, rs62508166, N_{eff}=211,793; signal 3, rs144239281, N_{eff}=208,431, in order of strength of association). Using 1000G Phase3 East Asian LD, variants are colored in red and blue with the first and second distinct signals

respectively (lead variants represented as diamonds). **(B)** Variant rs12549902, in high LD (EAS LD r^2=0.80, EUR r^2=0.83) with T2D signal 1, shows the strongest association with expression levels of *NXK6-3* in pancreatic islets in 118 individuals[16]. **(C)** Variant rs516946, in high LD (EAS LD r^2=0.96, EUR r^2=0.80) with T2D signal 2, shows the strongest association with expression levels of *ANK1* in subcutaneous adipose tissue in 770 individuals[19]. As rs62508166 is not available in the subcutaneous adipose tissue data set, a variant in perfect LD (rs28591316) was used and is represented by the blue diamond.

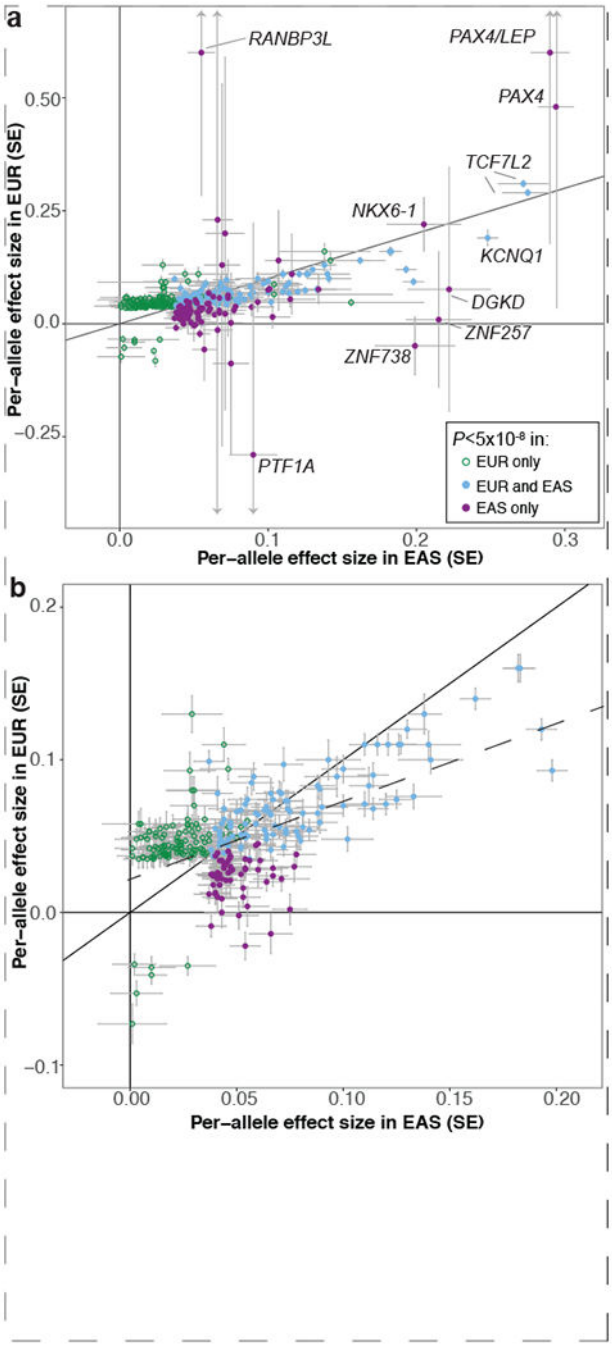


Figure 2: Effect size comparison of lead variants identified in this East Asian T2D GWAS BMI-unadjusted meta-analysis and previous European T2D GWAS meta-analysis[2].

For 332 unique lead variants identified from the two BMI unadjusted meta-analyses, per-allele effect sizes (β) from the European meta-analysis (*y*-axis) were plotted against per-allele effect sizes from this East Asian meta-analysis (*x*-axis). Effect sizes from both meta-analyses were from two-sided fixed-effect inverse-variance meta-analysis (maximal N_{eff}=211,793 for East Asian and 231,436 for European meta-analyses). Each point denotes the per-allele effect size; standard errors of the effect size estimates extend out as grey lines.

(A) All 332 lead variants; **(B)** 278 lead variants with minor allele frequency ≥5% in both ancestries. Variants are colored purple if they were significant ($P<5x10^{-8}$) in the East Asian analysis only, green if they were significant in European analysis only, and blue if they were significant in both the East Asian and European analyses (see Methods and Supplementary Table 7). The dashed diagonal line represents the trend line across all plotted variants. Compared to Supplementary Table 7, 70 variants are not plotted; 31 variants were present only in the analysis of East Asian individuals (median effect size 0.065; interquartile range 0.049-0.110) and 39 variants were present only in the analysis of European individuals (median effect size 0.083; interquartile range 0.063-0.170).

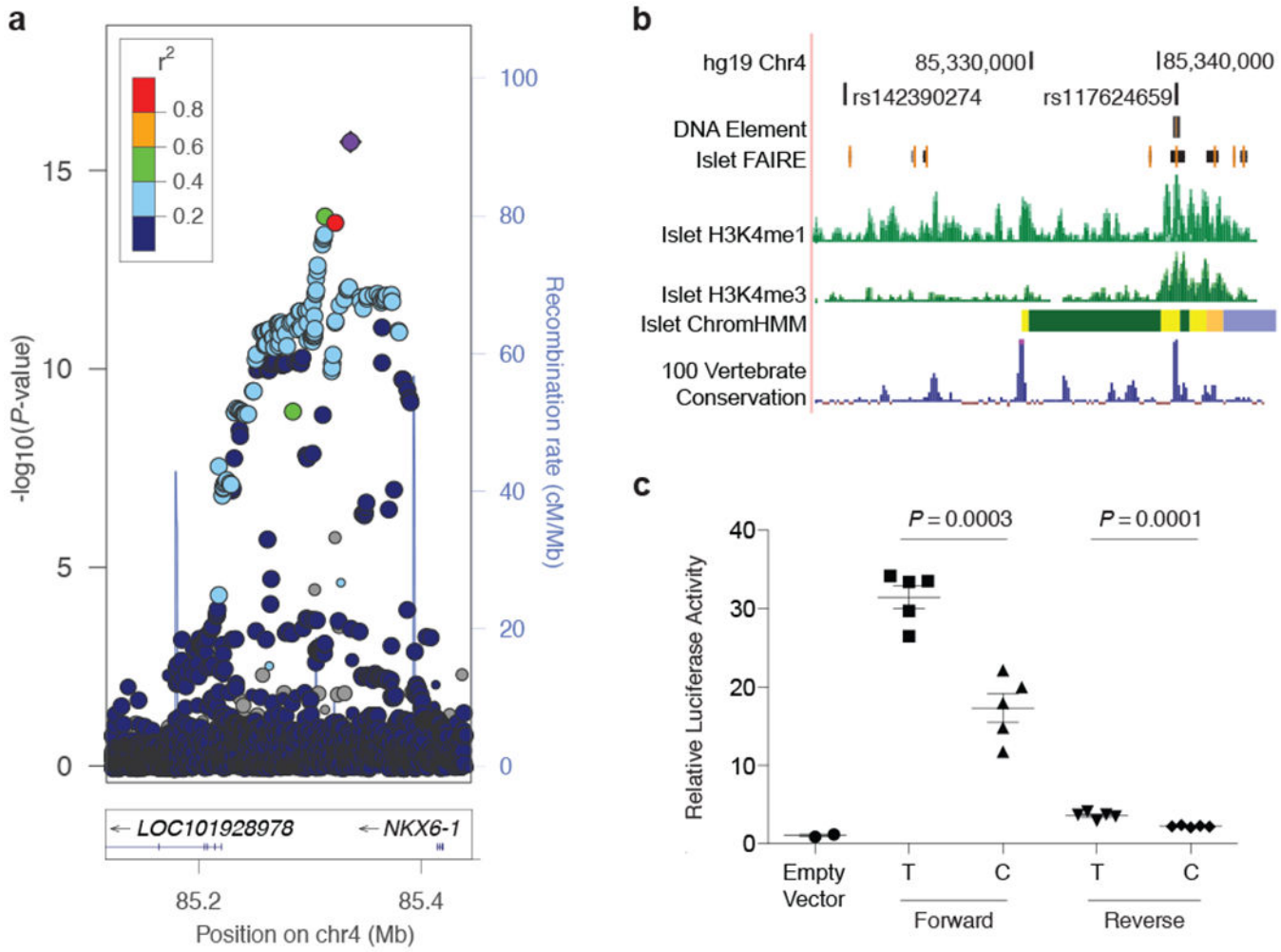


Figure 3: rs117624659 at *NKX6-1* locus exhibits allelic differences in transcriptional activity.

(A) rs117624659 (N_{eff}=211,214; purple diamond) shows the strongest association with T2D in the region. *P* values were from two-sided fixed-effect inverse-variance meta-analysis. Variants are colored based on 1000G Phase 3 East Asian LD with rs117624659. **(B)** rs117624659 and an additional candidate variant rs142390274 in high pairwise LD (r^2>0.80) span a 22 kb region approximately 75 kb upstream of *NKX6-1*. rs117624659 overlaps a region of open chromatin in pancreatic islets and lies within a region conserved across vertebrates. **(C)** rs117624659-T, associated with increased risk of T2D, showed greater transcriptional activity in an element cloned in both forward and reverse orientations with respect to *NKX6-1* in MIN6 cells compared to rs117624659-C and an "empty vector" containing a minimal promoter. Black lines represent mean (center horizontal line) and standard error (extended lines) relative luciferase activity from two-sided, unpaired t-tests using data from n=5 biologically independent samples/independent experiments.