

Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility

DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium^{1,2}, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium^{1,2}, South Asian Type 2 Diabetes (SAT2D) Consortium^{1,2}, Mexican American Type 2 Diabetes (MAT2D) Consortium^{1,2} & Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium^{1,2}

To further understanding of the genetic basis of type 2 diabetes (T2D) susceptibility, we aggregated published meta-analyses of genome-wide association studies (GWAS), including 26,488 cases and 83,964 controls of European, east Asian, south Asian and Mexican and Mexican American ancestry. We observed a significant excess in the directional consistency of T2D risk alleles across ancestry groups, even at SNPs demonstrating only weak evidence of association. By following up the strongest signals of association from the trans-ethnic meta-analysis in an additional 21,491 cases and 55,647 controls of European ancestry, we identified seven new T2D susceptibility loci. Furthermore, we observed considerable improvements in the fine-mapping resolution of common variant association signals at several T2D susceptibility loci. These observations highlight the benefits of trans-ethnic GWAS for the discovery and characterization of complex trait loci and emphasize an exciting opportunity to extend insight into the genetic architecture and pathogenesis of human diseases across populations of diverse ancestry.

The majority of GWAS of T2D susceptibility have been undertaken in populations of European ancestry¹⁻⁵, predominantly because of existing infrastructure, sample availability and relatively poor coverage by many of the earliest genome-wide genotyping arrays of common genetic variation in other major ethnic groups⁶. However, populations of European ancestry constitute only a subset of human genetic variation and are thus insufficient to fully characterize T2D risk variants in other ethnic groups. Furthermore, the latest genome-wide genotyping arrays are less biased toward Europeans, and more recent T2D GWAS have been performed with great success in populations from other ancestry groups, including east Asians⁷⁻¹², south Asians^{13,14}, Mexicans and Mexican Americans¹⁵ and African Americans¹⁶. These studies have provided initial evidence of overlap in T2D susceptibility loci between ancestry groups, as well as for coincident risk alleles at lead SNPs across diverse populations^{17,18}. These observations are consistent with a model in which the underlying causal variants at many of these loci are shared across ancestry groups and thus arose before migration of the human population out of Africa. Under such a model, we would expect to improve the power to detect new susceptibility loci for the disease and enhance the fine-mapping resolution of causal variants by combining GWAS across ancestry groups through trans-ethnic meta-analysis because of increased sample size and differences in the structure of linkage disequilibrium (LD) between such diverse populations^{6,19-21}.

In this study, we aggregated published meta-analyses of GWAS in a total of 26,488 cases and 83,964 controls from populations of European,

east Asian, south Asian and Mexican and Mexican American ancestry^{5,11,13,15}. T2D GWAS from populations of African ancestry, which would be expected to provide the greatest potential for fine mapping of common causal variants because of less extensive LD than other ethnic groups⁶, were not accessible for inclusion in our analyses. With these data, we aimed to (i) assess the evidence for excess concordance in the direction of effect of T2D risk alleles across ancestry groups; (ii) identify new T2D susceptibility loci through trans-ethnic meta-analysis and subsequent validation in an additional 21,491 cases and 55,647 controls of European ancestry; and (iii) evaluate the improvements in the fine-mapping resolution of common variant association signals in established T2D susceptibility loci through trans-ethnic meta-analysis despite the lack of GWAS from populations of African ancestry.

RESULTS

Study overview

We considered published meta-analyses of GWAS of T2D susceptibility from four major ethnic groups (**Supplementary Tables 1 and 2**) undertaken by (i) the DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium⁵ (European ancestry; 12,171 cases and 56,862 controls); (ii) the Asian Genetic Epidemiology Network T2D (AGEN-T2D) Consortium¹¹ (east Asian ancestry; 6,952 cases and 11,865 controls); (iii) the South Asian T2D (SAT2D) Consortium¹³ (south Asian ancestry; 5,561 cases and 14,458 controls); and (iv) the Mexican American T2D (MAT2D) Consortium¹⁵ (Mexican

Table 1 Concordance in the direction of effect of T2D risk alleles

European ancestry meta-analysis P value threshold	Trans-ethnic concordance								
	European into east Asian			European into south Asian			European into Mexican and Mexican American		
	Concordant SNPs/ total SNPs	%	Binomial test P	Concordant SNPs/ total SNPs	%	Binomial test P	Concordant SNPs/ total SNPs	%	Binomial test P
$P \leq 0.001$	180/316	57.0	0.0077	175/316	55.4	0.032	179/316	56.6	0.010
$0.001 < P \leq 0.01$	877/1,624	54.0	0.00068	861/1,624	53.0	0.0080	886/1,624	54.6	0.00013
$0.01 < P \leq 0.5$	2,556/5,053	50.6	0.21	2,604/5,053	51.5	0.015	2,588/5,053	51.2	0.043
$0.5 < P \leq 1$	2,535/5,039	50.3	0.34	2,532/5,039	50.2	0.37	2,519/5,039	50.0	0.51

The T2D risk alleles shown were identified in a meta-analysis of a GWAS of European ancestry (12,171 cases and 56,862 controls) along with those identified in meta-analyses of GWAS of east Asian (6,952 cases and 11,865 controls), south Asian (5,561 cases and 14,458 controls) and Mexican and Mexican American (1,804 cases and 779 controls) ancestry after exclusion of the 69 established autosomal susceptibility loci and defined as mapping within 500 kb of the previously reported lead SNP.

and Mexican American ancestry; 1,804 cases and 779 controls). We obtained association summary statistics from the four available ethnic-specific meta-analyses, each of which was imputed at up to 2.5 million autosomal SNPs from Phase II/III HapMap^{22,23}, to provide a uniform catalog of common genetic variation, defined by a minor allele frequency (MAF) of at least 5%, across ancestry groups (Online Methods). We then combined these association summary statistics across ancestry groups in a trans-ethnic fixed-effects meta-analysis (Online Methods).

Concordance of T2D risk alleles across ancestry groups

We began by evaluating heterogeneity in allelic effects (i.e., discordance in the direction and/or magnitude of odds ratios) between ancestry groups at 69 established autosomal T2D susceptibility loci. We assessed the evidence for heterogeneity at previously reported lead SNPs on the basis of Cochran's Q statistics from the trans-ethnic meta-analysis (Online Methods and **Supplementary Table 3**). We observed nominal evidence of heterogeneity (Bonferroni correction, $P_Q < 0.05/69 = 0.00072$) at the previously reported lead SNP at just three loci. At *TCF7L2* (rs7903146, $P_Q = 0.00055$), the odds ratio is the largest in populations of European ancestry, although the risk allele has a consistent direction of effect across ethnicities. At *PEPD* (rs3786897, $P_Q = 0.00055$) and *KLF14* (rs13233731, $P_Q = 0.00064$), however, the association signals are apparently specific to the populations of east Asian and European ancestry, respectively, despite the fact that the reported lead SNPs are common in all ethnic groups. We also observed that at 52 previously reported lead SNPs passing quality control in each of the four ethnic-specific meta-analyses, 34 showed the same direction of effect across all ancestry groups (65.4% compared with 12.5% expected by chance, binomial test $P < 2.2 \times 10^{-16}$). The strong evidence of homogeneity in allelic effects across ancestry groups at the majority of the previously reported lead SNPs argues against the 'synthetic association' hypothesis²⁴. It is improbable that GWAS signals at most of the established T2D susceptibility loci reflect unobserved lower-frequency causal alleles with larger effects because (i) rare variants are unlikely to have arisen before migration of the human population out of Africa and are thus not expected to be widely shared across diverse populations²⁵; and (ii) patterns of LD with these variants are anticipated to be highly variable between ethnicities.

To gain insights into the potential for the discovery of new T2D susceptibility loci through fixed-effects trans-ethnic meta-analysis, we next assessed the genome-wide coincidence of risk alleles (i.e., direction of effect) across ancestry groups after exclusion of the 69 established autosomal GWAS signals, which we defined as mapping within 500 kb of the previously reported lead SNPs (Online Methods). First we identified independent SNPs (separated by at least 500 kb) with nominal evidence of association ($P \leq 0.001$) with T2D

from the European ancestry meta-analysis. By aligning the effect of the T2D risk allele from the European meta-analysis into the other ancestry groups, we observed evidence of a significant excess in directional concordance between ethnicities: 57.0% with east Asian populations (binomial test $P = 0.0077$); 55.4% with south Asian populations (binomial test $P = 0.032$); and 56.6% with Mexican and Mexican American populations (binomial test $P = 0.010$). Using the same approach, we also observed an excess of consistency in the direction of effect between ethnicities at independent SNPs demonstrating weaker evidence of T2D association ($0.001 < P \leq 0.01$) from the European meta-analysis (**Table 1**). In contrast, when we considered independent SNPs with no evidence of association ($P > 0.5$) with T2D, there was no enrichment in coincident risk alleles across ethnic groups. We repeated this analysis by identifying T2D risk alleles at SNPs with nominal evidence of association in each of the east Asian, south Asian and Mexican and Mexican American meta-analyses and assessing concordance in the direction of effect in each of the other ancestry groups (**Supplementary Table 4**). The evidence for an excess in concordance between T2D risk alleles across ethnicities was not as strong, particularly for the Mexican and Mexican American meta-analysis. However, this finding presumably reflects reduced power due to smaller sample sizes, and there was still significant over-representation of alleles with the same direction of effect across ancestry groups at SNPs with nominal evidence of association with the disease.

Seven new T2D susceptibility loci at genome-wide significance

The observations from our concordance analyses are consistent with a long tail of common T2D susceptibility variants with effects that are decreasing in magnitude but that are homogeneous across ancestry groups. Under such a model, we would expect these variants to be amenable to discovery by trans-ethnic fixed-effects meta-analyses. In this study, by aggregating the published ethnic-specific meta-analyses under a fixed-effects model, we identified 33 independent SNPs (separated by at least 500 kb) with suggestive evidence of association ($P < 10^{-5}$) at loci not previously reported for T2D susceptibility in any ancestry group (**Supplementary Table 5** and **Supplementary Fig. 1**). By convention, we labeled loci according to the gene nearest to the lead SNP unless a compelling biological candidate mapped nearby. It is essential to validate partially imputed association signals with direct genotyping. Consequently, we carried forward these 33 loci for *in silico* follow up in a meta-analysis of an additional 21,491 cases and 55,647 controls of European ancestry⁵ genotyped with the MetaboChip (Online Methods and **Supplementary Tables 1** and **2**). This custom array was designed to facilitate cost-effective replication of nominal associations for T2D and other metabolic and cardiovascular traits²⁶. However, it provides relatively limited coverage of common genetic variation genome wide, with the result that the lead

Table 2 New T2D susceptibility loci achieving genome-wide significance ($P < 5 \times 10^{-8}$)

Locus	Lead SNP	Chr	Build 36		Trans-ethnic discovery		European ancestry validation		Combined meta-analysis		
			position (bp)	Risk	Other	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
<i>TMEM154</i>	rs6813195	4	153,739,925	C	T	1.08 (1.05–1.11)	4.2×10^{-9}	1.08 (1.05–1.11)	2.0×10^{-6}	1.08 (1.06–1.10)	4.1×10^{-14}
<i>SSRI-RREB1</i>	rs9505118	6	7,235,436	A	G	1.06 (1.04–1.09)	1.9×10^{-6}	1.06 (1.03–1.09)	1.7×10^{-4}	1.06 (1.04–1.08)	1.4×10^{-9}
<i>FAF1</i>	rs17106184	1	50,682,573	G	A	1.11 (1.07–1.16)	1.9×10^{-6}	1.09 (1.04–1.15)	4.8×10^{-4}	1.10 (1.07–1.14)	4.1×10^{-9}
<i>POU5F1-TCF19</i>	rs3130501	6	31,244,432	G	A	1.07 (1.04–1.10)	1.5×10^{-6}	1.06 (1.03–1.10)	7.0×10^{-4}	1.07 (1.04–1.09)	4.2×10^{-9}
<i>LPP</i>	rs6808574	3	189,223,217	C	T	1.08 (1.04–1.11)	4.3×10^{-6}	1.06 (1.03–1.09)	2.6×10^{-4}	1.07 (1.04–1.09)	5.8×10^{-9}
<i>ARL15</i>	rs702634	5	53,307,177	A	G	1.08 (1.05–1.11)	3.4×10^{-7}	1.05 (1.02–1.08)	2.1×10^{-3}	1.06 (1.04–1.09)	6.9×10^{-9}
<i>MPHOSPH9</i>	rs4275659	12	122,013,881	C	T	1.06 (1.03–1.09)	5.5×10^{-6}	1.06 (1.02–1.09)	4.4×10^{-4}	1.06 (1.04–1.08)	9.5×10^{-9}

The T2D susceptibility loci shown were identified through trans-ethnic discovery-stage GWAS meta-analysis of 26,488 cases and 83,964 controls of European, east Asian, south Asian and Mexican and Mexican American ancestry with follow up in a validation-stage meta-analysis of an additional 21,491 cases and 55,647 controls of European ancestry genotyped with the Metabochip. Chr, chromosome; OR, odds ratio; CI, confidence interval.

^aAlleles are aligned to the forward strand of NCBI Build 36.

SNPs, or close proxies (CEU $r^2 > 0.6$ from Phase II HapMap), were present at just 24 of the loci. We also identified poorer proxies at two additional loci, rs9505118 (*SSRI-RREB1*, CEU $r^2 = 0.26$, $P = 1.9 \times 10^{-6}$) and rs4275659 (*MPHOSPH9*, CEU $r^2 = 0.48$, $P = 5.5 \times 10^{-6}$), which nonetheless demonstrated only marginally weaker association signals than the lead SNPs (*SSRI-RREB1*, rs9502570, $P = 5.7 \times 10^{-7}$; *MPHOSPH9*, rs1727313, $P = 1.2 \times 10^{-6}$). Given that these variants met our threshold for follow up from the trans-ethnic meta-analysis, we also considered them for validation.

By combining association summary statistics from the trans-ethnic discovery and European ancestry validation meta-analyses, SNPs at seven loci achieved genome-wide significance (combined meta-analysis $P < 5 \times 10^{-8}$) (Table 2 and Fig. 1). We observed no evidence of heterogeneity in allelic effects between the discovery and validation stages of the combined meta-analysis (Supplementary Table 5). As we expected, the new loci are characterized by lead SNPs that are relatively common in all ethnicities and have modest effects on T2D susceptibility that are homogeneous across ancestry groups (Supplementary Table 6). We did not harmonize adjustments for covariates within or between consortia because of variation in individual study design and recorded non-genetic risk factors. However, we observed no evidence of heterogeneity in allelic effects in the European ancestry validation meta-analysis after stratification of the studies according to covariate adjustment (Online Methods and Supplementary Table 7). These data thus provide no evidence of bias in the allelic effect estimates at lead SNPs at the new loci and suggest our results to be robust to variability in correction for potential confounders across studies.

The new loci include SNPs mapping near *POU5F1-TCF19* in the major histocompatibility complex (MHC), a region of the genome that is essential to the immune response. The MHC harbors HLA class II genes, which together account for approximately half of the genetic risk to type 1 diabetes (T1D)²⁷. We observed no evidence of association of T2D with tags for traditional T1D HLA risk alleles in the trans-ethnic meta-analysis: *HLA-DR4* (rs660895, $P = 0.32$) and *HLA-DR3* (rs2187668, $P = 0.34$). Furthermore, when we considered lead SNPs at 49 T1D susceptibility loci (Supplementary Table 8), we observed nominal evidence of association ($P < 0.05$) with T2D and the same risk allele for both diseases at just two loci (*GLIS3* and 6q22.32) but not at that mapping to the MHC (rs9268645, $P = 0.33$). There is very strong evidence that T1D risk variants, particularly in the MHC, are also associated with latent autoimmune diabetes of adulthood (LADA)^{28,29}, which is a late age-onset, more indolent form of the disease that often results in a clinical misdiagnosis of T2D. Although studies contributing to the trans-ethnic meta-analysis

differed in the degree to which they were able to exclude cases with LADA, the lack of association of T1D risk variants suggests that the rates of diagnostic misclassification of autoimmune diabetes were too modest to drive the T2D GWAS signal at the *POU5F1-TCF19* locus.

The new loci also include SNPs mapping to *ARL15* and *SSRI-RREB1*, which have been implicated previously at genome-wide significance in the regulation of fasting insulin (FI) and fasting glucose (FG), respectively³⁰. The lead SNPs for T2D (rs702634) and FI (rs4865796) mapping to *ARL15* are closely correlated in populations of European and east Asian ancestry (CEU $r^2 = 1.00$ and CHB+JPT $r^2 = 0.87$ from Phase II HapMap). However, the lead T2D SNP (rs9505118) is independent of that for FG (rs17762454) at the *SSRI-RREB1* locus (CEU and CHB+JPT $r^2 < 0.05$). The *ARL15* locus has also been associated with circulating levels of adiponectin, which is an adipocyte-secreted protein that has antidiabetic effects³¹, but the lead SNP (rs4311394) is independent of that for T2D susceptibility from the trans-ethnic meta-analysis.

To obtain a more comprehensive view of the overlap of the newly associated T2D susceptibility loci with metabolic phenotypes, we interrogated published European ancestry meta-analyses from the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC) Investigators^{3,30}, the Genetic Investigation of Anthropometric Traits (GIANT) Consortium^{32,33} and the Global Lipids Genetics Consortium³⁴ to evaluate the effect of T2D risk alleles on glycemic traits, including homeostatic model of assessment indices of beta-cell function (HOMA-B) and insulin resistance (HOMA-IR); anthropometric measures; and plasma lipid concentrations (Online Methods and Supplementary Tables 9–11). T2D risk alleles at *SSRI-RREB1* and *LPP* have features that indicate a primary role in susceptibility through beta-cell dysfunction: increased FG ($P = 1.0 \times 10^{-5}$ and $P = 8.6 \times 10^{-7}$, respectively) and reduced HOMA-B ($P = 0.11$ and $P = 0.011$, respectively). Conversely, the T2D risk allele mapping to *ARL15* is associated with increased FI, most strongly after adjustment for body mass index (BMI) ($P = 5.0 \times 10^{-12}$), and increased HOMA-IR ($P = 0.021$) and is thus more characteristic of action through insulin resistance. This risk allele is also associated with reduced levels of high-density lipoprotein cholesterol ($P = 0.022$) and increased levels of triglycerides ($P = 0.010$), as we expected, but also with reduced BMI ($P = 5.6 \times 10^{-5}$).

To identify the most promising functional candidate transcripts among those mapping to the new susceptibility loci, we interrogated public databases and unpublished resources for expression quantitative trait loci (eQTL) from a variety of tissues (Online Methods). The lead T2D SNPs at three loci showed nominal association ($P < 10^{-5}$)

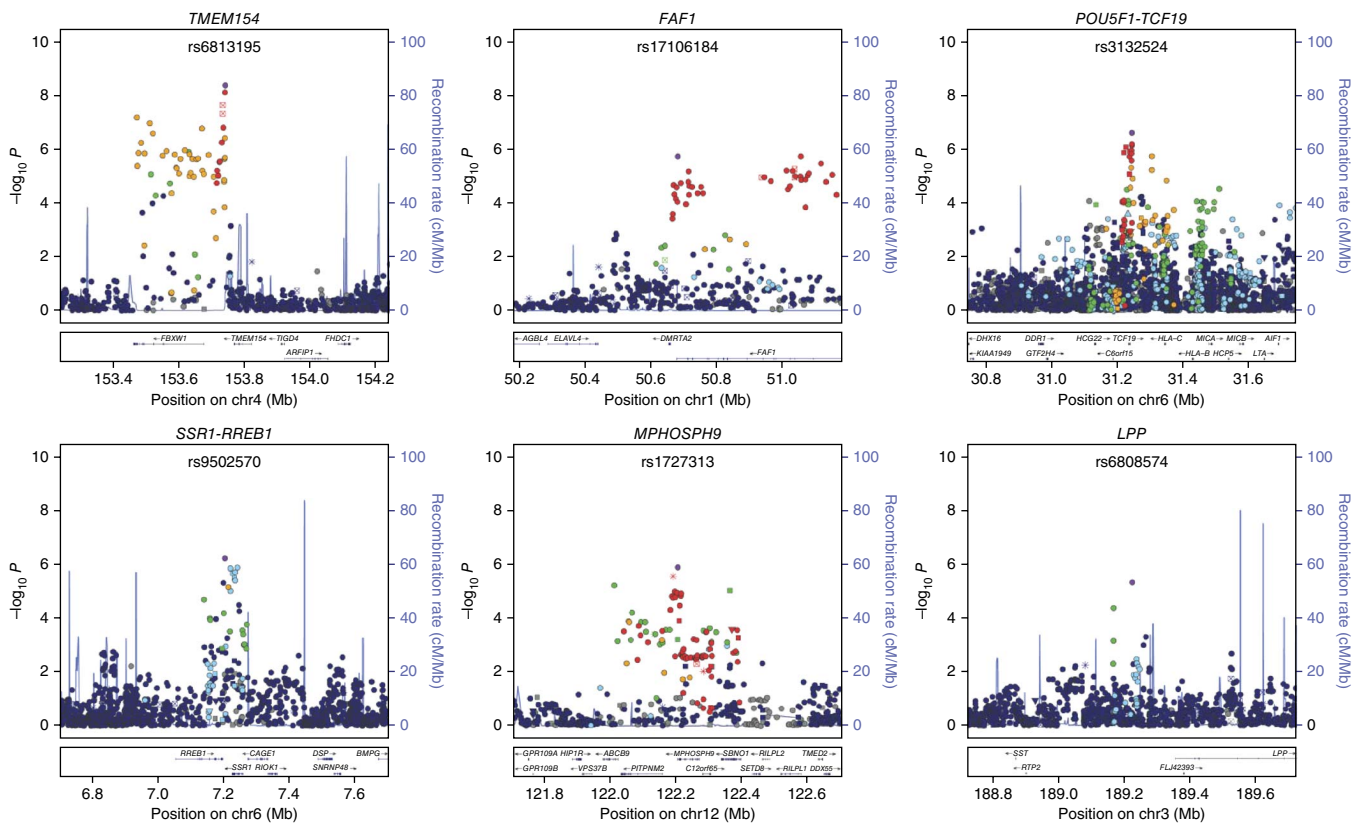


Figure 1 Signal plots of the trans-ethnic discovery-stage GWAS meta-analysis for new T2D susceptibility loci. The trans-ethnic meta-analysis comprised 26,488 T2D cases and 83,964 controls from populations of European, east Asian, south Asian and Mexican and Mexican American ancestry imputed up to 2.5 million Phase II/III HapMap autosomal SNPs. Each point represents a SNP passing quality control in the trans-ethnic meta-analysis plotted with its P value (on a $-\log_{10}$ scale) as a function of genomic position (NCBI Build 36). In each plot, the lead SNP is represented by the purple symbol. The color coding of all other SNPs indicates LD with the lead SNP (estimated by CEU r^2 values from Phase II HapMap): red, $r^2 \geq 0.8$; gold, $0.6 \leq r^2 < 0.8$; green, $0.4 \leq r^2 < 0.6$; cyan, $0.2 \leq r^2 < 0.4$; blue, $r^2 < 0.2$; gray, r^2 unknown. The shape of the plotting symbol corresponds to the annotation of the SNP: upward triangle for frameshift or splice; downward triangle for nonsynonymous; square for synonymous or UTR; and circle for intronic or noncoding. Recombination rates are estimated from Phase II HapMap, and gene annotations are taken from the UCSC genome browser.

with expression and were in strong LD (CEU and CHB+JPT $r^2 > 0.8$) with the reported *cis*-eQTL variant: *SSR1* (B cells, $P = 2.2 \times 10^{-6}$) at the *SSR1-RREB1* locus; *ABC9* (liver, $P = 7.4 \times 10^{-12}$) and *SETD8* (lung, $P < 2.0 \times 10^{-16}$) at the *MPHOSPH9* locus; and *HCG27* (monocytes, $P = 1.3 \times 10^{-69}$) at the *POU5F1-TCF19* locus (**Supplementary Table 12**).

We also evaluated new loci for potential functional mechanisms underlying T2D susceptibility (Online Methods). We identified variants for functional annotation in pilot data from the 1000 Genomes Project²⁵ that are in strong LD (CEU and CHB+JPT $r^2 > 0.8$) with the lead SNPs in the seven new susceptibility loci. We identified a missense variant at the *POU5F1-TCF19* locus in *TCF19* (rs113581344, p.Val211Met; CEU $r^2 = 0.96$ and CHB+JPT $r^2 = 0.80$ with lead SNP rs3130501), although it is predicted to be tolerated by SIFT³⁵ (**Supplementary Table 13**). Lead SNPs in the new susceptibility loci were also in strong LD with variants in the UTRs of *SSR1* (at the *SSR1-RREB1* locus) and *ABC9*, *OGFOD2* and *PITPNM2* (at the *MPHOSPH9* locus). Variants in strong LD with the lead SNPs at two

of the new susceptibility loci overlap regions of predicted regulatory function generated by the ENCODE Project³⁶ (**Supplementary Fig. 2**). The lead SNP at the *LPP* locus maps to an enhancer region that is active in HepG2 cells. We also identified a variant at the *FAF1* locus (rs58836765; CEU $r^2 = 0.89$ and CHB+JPT $r^2 = 0.80$ with lead SNP rs17106184) that overlaps a region of open chromatin activity in pancreatic islets and other cell types. This open chromatin site is in a region that is correlated with expression of *ELAVL4*, which has been demonstrated to regulate insulin translation in pancreatic beta cells³⁷, highlighting this transcript as a credible candidate at the *FAF1* locus. Regulatory annotations in HepG2 cells and pancreatic islets are both broadly enriched at T2D-associated variants³⁸ and are thus supportive of these functional mechanisms for causal variant activity at both loci.

Improved fine-mapping resolution at T2D susceptibility loci

Given our observation that the causal variants underlying GWAS signals are shared across ancestry groups at many T2D susceptibility

Table 3 Properties of the 99% credible sets of SNPs at ten established T2D susceptibility loci

Locus	Chr	99% credible set: European ancestry meta-analysis			99% credible set: trans-ethnic meta-analysis			99% credible set: reduction	
		SNPs	Interval (bp)	Build 36 location (bp)	SNPs	Interval (bp)	Build 36 location (bp)	SNPs	Interval (bp)
<i>JAZF1</i>	7	9	75,685	28,147,081–28,222,765	4	15,667	28,147,081–28,162,747	5	60,018
<i>SLC30A8</i>	8	4	35,488	118,253,964–118,289,451	2	243	118,253,964–118,254,206	2	35,245
<i>CDKAL1</i>	6	5	24,244	20,787,688–20,811,931	2	1,549	20,794,552–20,796,100	3	22,695
<i>HHEX/IDE</i>	10	8	19,195	94,452,862–94,472,056	2	937	94,455,539–94,456,475	6	18,258
<i>TCF7L2</i>	10	3	13,684	114,744,078–114,757,761	2	2,309	114,746,031–114,748,339	1	11,375
<i>IGF2BP2</i>	3	17	32,656	186,980,329–187,012,984	12	24,504	186,988,481–187,012,984	5	8,152
<i>FTO</i>	16	27	45,981	52,357,008–52,402,988	10	39,335	52,361,075–52,400,409	17	6,646
<i>CDKN2A/B</i>	9	3	2,019	22,122,076–22,124,094	1	1	22,122,076–22,122,076	2	2,018
<i>PPARG</i>	3	23	265,269	12,106,687–12,371,955	21	265,269	12,106,687–12,371,955	2	0
<i>MTNR1B</i>	11	15	55,032	92,307,378–92,362,409	15	55,032	92,307,378–92,362,409	0	0

The properties shown are based on association summary statistics from the meta-analysis of the European ancestry GWAS only (12,171 cases and 56,862 controls) and the trans-ethnic meta-analysis of European, east Asian, south Asian and Mexican and Mexican American ancestry GWAS (26,488 cases and 83,964 controls).

loci, we evaluated the evidence for improved fine-mapping resolution through trans-ethnic meta-analysis. For this purpose, we combined association summary statistics from the ethnic-specific meta-analyses using MANTRA³⁹. This Bayesian approach has the advantage of allowing for heterogeneity in allelic odds ratios between ancestry groups arising as a result of differential patterns of LD with a shared underlying causal variant across diverse populations, which cannot be accommodated in fixed-effects meta-analysis (Online Methods). Simulation studies have demonstrated improved detection and localization of causal variants through trans-ethnic meta-analysis with MANTRA compared to either a fixed- or random-effects model^{39,40}.

Within each locus, we constructed credible sets⁴¹ of SNPs that are most likely to be causal on the basis of their statistical evidence of association from the MANTRA meta-analysis. Credible sets can be interpreted in a way similar to confidence intervals in a frequentist statistical framework. For example, assuming that a locus harbors a single causal variant that is reported in the meta-analysis, the probability that it will be contained in the 99% credible set is 0.99. Smaller credible sets, in terms of the number of SNPs they contain or the genomic interval they cover, thus correspond to fine mapping at higher resolution. It is essential that SNP coverage be as uniform as possible across studies in the construction of credible sets, otherwise differences in association signals between variants may reflect variability in sample sizes in the meta-analysis and not true differences in the magnitude of effects on T2D susceptibility. Consequently, we did not consider the European ancestry Metabochip validation studies in our fine-mapping analyses because the SNP density across the majority of T2D susceptibility loci on the array is too sparse to allow high-quality imputation up to the Phase II/III HapMap reference panels utilized in the trans-ethnic discovery GWAS.

In constructing credible sets, we assume that there is a single causal variant at each locus. However, there is increasing evidence that multiple association signals, typically characterized by independent common index SNPs, are relatively widespread at T2D susceptibility loci, for example *CDKN2A-CDKN2B* and *KCNQ1* (ref. 6). Fine mapping of these independent association signals will require formal conditioning, adjusting for genotypes at each index SNP in turn, before construction of the credible set for each underlying causal variant. Approximate conditioning, without formal computation, as implemented in GCTA⁴² makes use of meta-analysis summary statistics and a reference panel to approximate LD between SNPs (and hence correlation between parameter estimates in a joint association model). Unfortunately, this approach is not feasible in a trans-ethnic context because of differences in LD structure between ancestry groups and thus could not be applied in this study. Consequently, the credible sets

defined here correspond to fine mapping across association signals at each locus.

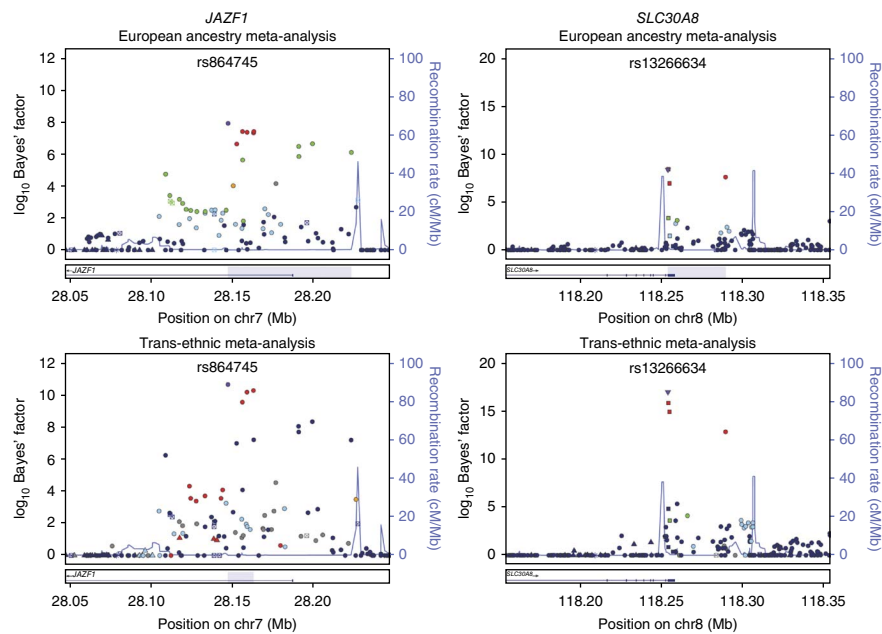
To assess the improvements in fine-mapping resolution by combining GWAS from diverse populations, we compared the properties of the MANTRA 99% credible set on the basis of association summary statistics from the European ancestry-only meta-analysis and the trans-ethnic meta-analysis of European, east Asian, south Asian and Mexican and Mexican American ancestry groups. We focused on 10 autosomal loci (of the 69 previously established) that attained association with T2D susceptibility at genome-wide significance in the European ancestry meta-analysis (Table 3). We did not consider loci with weaker signals of association, as they were typically characterized by large 99% credible sets in the European ancestry meta-analysis and thus might provide an overestimate of the improvement in fine-mapping resolution by combining GWAS across ancestry groups. Of the loci considered, only *MTNR1B* did we not see any improvement in fine-mapping resolution in terms of the number of SNPs and the genomic interval covered by the 99% credible set after trans-ethnic meta-analysis.

We observed the greatest enhancement in fine-mapping resolution after trans-ethnic meta-analysis at the *JAZF1* locus, where the genomic interval covered by the 99% credible set was reduced from 76 kb to just 16 kb (Fig. 2 and Supplementary Fig. 3). Of the nine variants in the European 99% credible set, five were excluded after trans-ethnic meta-analysis because of low LD with the lead SNP at this locus in populations of east Asian ancestry (CHB+JPT $r^2 < 0.05$ with rs864745). Among the variants retained in the 99% credible set after trans-ethnic meta-analysis, interrogation of predicted regulatory function from the ENCODE Project³⁶ showed that rs1635852 maps to a region of open chromatin with enhancer activity that is bound by several transcription factors. This SNP has been shown previously to have allelic differences in pancreatic islet enhancer activity⁴³ and is also correlated with expression of *CREB5*, highlighting this transcript as a credible candidate at the *JAZF1* locus.

We also observed a substantial reduction in the genomic interval covered by the credible set at the *SLC30A8* locus (Fig. 2 and Supplementary Fig. 3) from 35 kb (four SNPs) on the basis of the European ancestry GWAS only to less than 1 kb (two SNPs) after trans-ethnic meta-analysis. However, the lead SNP is strongly correlated with all variants in the credible set before trans-ethnic meta-analysis in both the European and east Asian ancestry groups (CEU and CHB+JPT $r^2 \geq 0.8$ with rs13266634), suggesting that the improved fine-mapping resolution at this locus is more likely the result of increased sample size than differences in LD structure between the populations. Encouragingly, the lead SNP after trans-ethnic meta-analysis is more clearly separated from the other SNPs

Figure 2 Signal plots presenting 99% credible sets of SNPs at the *JAZF1* and *SLC30A8* loci.

The credible sets were constructed on the basis of the meta-analysis of European ancestry GWAS only (12,171 cases and 56,862 controls) and the trans-ethnic meta-analysis of European, east Asian, south Asian and Mexican and Mexican American ancestry GWAS (26,488 cases and 83,964 controls). In each plot, each point represents a SNP passing quality control in the MANTRA analysis plotted with its Bayes' factor (on a \log_{10} scale) as a function of genomic position (NCBI Build 36). The lead SNP is represented by the purple symbol. The color coding of all other SNPs indicates LD with the lead SNP (estimated by Phase II HapMap CEU r^2 values for the European ancestry meta-analysis and CHB+JPT for the trans-ethnic meta-analysis to highlight differences in structure between ancestry groups): red, $r^2 \geq 0.8$; gold, $0.6 \leq r^2 < 0.8$; green, $0.4 \leq r^2 < 0.6$; cyan, $0.2 \leq r^2 < 0.4$; blue, $r^2 < 0.2$; gray, r^2 unknown. The shape of the plotting symbol corresponds to the annotation of the SNP: upward triangle for frameshift or splice; downward triangle for nonsynonymous; square for synonymous or UTR; and circle for intronic or noncoding. Recombination rates are estimated from Phase II HapMap, and gene annotations are taken from the UCSC genome browser. The genomic region covered by the 99% credible set is highlighted in gray.



in the credible set and is a nonsynonymous variant, p.Arg325Trp, that has an established functional role in T2D susceptibility⁴⁴.

We next tested variants present in the 99% credible sets at the ten loci on the basis of the European ancestry GWAS only and the trans-ethnic meta-analysis for enrichment of functional annotation compared to randomly shifted element locations (Online Methods). Variants in the trans-ethnic 99% credible sets were significantly enriched (empirical $P < 0.05$) for overlap with DNaseI hypersensitive sites (DHS $P = 0.038$) and transcription factor binding sites (TFBS $P = 0.0060$). However, we observed no such enrichment in either annotation category for the European ancestry 99% credible sets (DHS $P = 0.18$; TFBS $P = 0.087$). These data suggest that variants retained after trans-ethnic meta-analysis show greater potential for functional impact on T2D susceptibility through these regulatory mechanisms.

The fine-mapping intervals defined by credible sets after trans-ethnic meta-analysis are limited by the density and allele frequency spectrum of the GWAS genotyping arrays and HapMap reference panels used for imputation. Although these reference panels provide comprehensive coverage of common SNPs (MAF $> 5\%$) across ancestry groups, imputation up to phased haplotypes from the 1000 Genomes Project^{25,45}, for example, would allow assessment of the impact of lower-frequency variation on T2D susceptibility in diverse populations^{46–48}. However, we have demonstrated that for a fixed reference panel, trans-ethnic meta-analysis can improve the localization of common causal SNPs within established T2D susceptibility loci, and we have identified highly annotated variants within fine-mapping intervals defined by the 99% credible sets. We have also assessed the sensitivity of the trans-ethnic fine-mapping analysis to genotype quality at directly typed or imputed SNPs (**Supplementary Table 14**). We repeated MANTRA fine mapping with subsets of SNPs that passed quality control in at least 80% ($n = 88,361$) or 90% ($n = 99,406$) of individuals from the trans-ethnic meta-analysis. As the threshold for the reported sample size increased, the number of SNPs included in the fine-mapping analysis was reduced, but

the genomic intervals covered by the 99% credible sets remained unchanged, suggesting the resolution to be relatively robust to genotype quality at common variants.

DISCUSSION

We have identified seven new loci for T2D susceptibility at genome-wide significance by combining GWAS from multiple ancestry groups. Our study has provided evidence of many more common variant loci not yet reaching genome-wide significance that contribute to the heritability of T2D susceptibility, which is in agreement with polygenic analyses in European ancestry GWAS^{5,49}. The effects of these common variants are modest but are homogeneous across ancestry groups and would thus be amenable to discovery through trans-ethnic meta-analysis in larger samples. We have also demonstrated improvements in the resolution of fine mapping of common variant association signals through trans-ethnic meta-analysis, even in the absence of GWAS of African ancestry, which would be expected to better refine localization because of reduced LD in these populations. Future releases of reference panels from the 1000 Genomes Project are anticipated to comprise 2,500 samples, including haplotypes of south Asian ancestry and a wider representation of populations of African descent. This panel will provide a comprehensive catalog of genetic variation with MAFs as low as 0.5%, as well as many rarer variants, across major ancestry groups, thus facilitating imputation and coverage of loci for future trans-ethnic fine-mapping efforts.

Our analyses clearly highlight the benefits of combining GWAS from multiple ancestry groups for discovery and characterization of common variant loci contributing to complex traits and emphasize an exciting opportunity to further our understanding of the biological mechanisms underlying human diseases across populations from diverse ethnicities.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

ACKNOWLEDGMENTS

Funding for the research undertaken in this study has been received from the following: the Canadian Institutes of Health Research; the European Commission (ENGAGE FP7 HEALTH-F4-2007- 201413); the Medical Research Council UK (G0601261); the Mexico Convocatoria (SSA/IMMS/ISSSTE-CONACYT 2012-2, clave 150352, IMSS R-2011-785-018 and CONACYT Salud-2007-C01-71068); the US National Institutes of Health (DK062370, HG000376, DK085584, DK085545, DK073541 and DK085501); and the Wellcome Trust (WT098017, WT090532, WT090367, WT098381, WT081682 and WT085475). We acknowledge the many colleagues who contributed to collection and phenotypic characterization of the clinical samples and the genotyping and analysis of the GWAS data, full details of which are provided in the contributing consortia papers^{5,11,13,15}. We also thank those individuals who agreed to participate in this study.

AUTHOR CONTRIBUTIONS

Writing group: A. Mahajan, M.J.G., W. Zhang, J.E.B., K.J.G., M.H., A.D.J., I.P., E.Z., Y.Y.T., M.B., E.J.P., J.C.C., E.S.T., M.I.M. and A.P.M.

Analysis group: A. Mahajan, M.J.G., W. Zhang, J.E.B., K.J.G., T. Ferreira, M.H., A.D.J., M.C.Y.N., I.P., D.S., X.W., E.Z., Y.Y.T., M.B., E.J.P., M.I.M. and A.P.M.

DIAGRAM Consortium samples, genotyping, analysis and management:

A. Mahajan, M.H., I.P., D.S., E.Z., G.R.A., P.A., M.A., D.B., B.B., I.B., J.B., R.B., R.N.B., B.O.B., E.B., L.L.B., N.B., H. Campbell, J.C., S.C., G.C., H. Chen, P.S.C., F.S.C., M.C.C., D.J.C., A.T.C., R.M.v.D., J. Danesh, U.d.F., G.D., P.D., A.S.D., C.D., A.S.F.D., P.J.D., M.D., C.v.D., J. Dupuis, S.E., V.E., R.E., J.G.E., T.E., E.E., T. Ferreira, J.C.F., P. Fontanillas, N.G.F., T. Forsen, C.F., R.M.F., T.M.F., P. Froguel, K.G., C. Gieger, B.G., H.G., G.B.G., L.C.G., C.J.G., C. Guiducci, A. Hamsten, A.T.H., C. Hayward, C. Herder, A. Hofman, O.L.H., K. Hovingh, A.B.H., F.B.H., J.H., S.E. Humphries, S.E. Hunt, D.J.H., K. Hveem, T.I., E.I., B.I., A.U.J., A. James, K.-H.J., A. Jonsson, H.M.K., S. Kanoni, W.H.L.K., S. Kathiresan, S.M.K.-K., H.K., K.-T.K., L.K., N. Klopp, A. Kong, E.K.-H., P. Kraft, J. Kravic, A. Kumar, J. Kuusisto, M. Laakso, V. Lagou, T.A.L., C. Langenberg, C. Langford, R.L., K.L., M. Li, L.L., C.M.L., E.L., C.-T.L., S. Lobbens, R.J.F.L., J. Luan, V. Lyssenko, R.M., S. Männistö, J.B.M., O.M., A. Metspalu, J.M., G.M., E.M., S. Moebus, K.L.M., A.D.M., T.W.M., M.M.-N., B.M., P.N., P.M.N., I.N., M.M.N., K.R.O., C.N.A.P., J.S.P., M.P., S. Pechlivanis, N.L.P., L.P., J.R.B.P., A.P., C.G.P.P., S. Potter, J.F.P., L.Q., L.R., W.R., R.R., S. Raychaudhuri, N.W.R., E.R., S. Ripatti, N.R., M.R., E.J.R., I.R., D.R., T.E.S., V. Salomaa, J. Saltevo, J. Saramies, L.J.S., R.A.S., A.V.S., B.S., S. Shah, A.R.S., G. Sigurdsson, E.S., A. Silveira, S. Sivapalaratnam, A. Stančáková, K. Stefansson, G. Steinbach, V. Steinthorsdottir, K. Stirrups, R.J.S., H.M.S., Q.S., A.-C.S., T.M.T., B.T., G.T., U.T., E. Tikkanen, J. Trakalo, E. Tremoli, M.D.T., T.T., J. Tuomilehto, A.G.U., S.V., F.V., B.F.V., N.J.W., R.W., T.W., J.F.W., S.W., W.W., A.R. Wood, L.Y., D.Z., D.A., M.B., M.I.M. and A.P.M.

AGEN-T2D Consortium samples, genotyping, analysis and management:

M.J.G., X.W., L.S.A., T.A., Y.B., Q.C., J.C.N.C., L.-C.C., T.-J.C., Y.-C.C., C.-H.C., Y.-T.C., N.H.C., Y.M.C., L.-M.C., Y.G., B.-G.H., K. Hara, A.K.H., C. Hu, F.B.H., H.I., W.J., T.K., N. Kato, H.-L.K., S. Kim, Y.J.K., S.H.K., J.-M.L., N.R.L., Y.L., J.J.L., J. Long, W.L., R.C.W.M., S. Maeda, K.L.M., J.N., E.N., P.-K.N., K.O., T.H.O., K.S.P., X.O.S., X.S., W.Y.S., R.T., W.T.T., F.J.T., C.W., T.Y.W., J.-Y.W., Y.W., K.Y., T.Y., M. Yokota, R.Z., W. Zheng, Y.S.C., J.-Y.L., M. Seielstad, Y.Y.T., E.S.T. and M.I.M.

SAT2D Consortium samples, genotyping, analysis and management:

W. Zhang, I.P., D.S., G.R.A., T.A., A.H.B., A.B., L.F.B., M. Caulfield, K.-S.C., M. Chidambaram, J. Danesh, D.D., P.D., A.S.D., P.E., T.M.F., P. Froguel, P. Frossard, E.G., N.H., A.K.H., Z.I.H., M.I., T.J., J.B.M.J., N. Kato, P. Katulanda, A.M.K., C.-C.K., S. Kowlessur, M.M.K., X.L., J. Liang, S. Liju, W.-Y.L., J.J.L., D.R.M., V.M., A.C.N., J.M.P., V.R., A.R., S.D.R., M. Samuel, D.K.S., J. Scott, J. Sehmi, N.S., A.S.S., X.S., K.S.S., C.S., R.T., F.T., A.R. Wickremasinghe, T.Y.W., M. Yang, R.Y., F.Z., P.Z.Z., J. Kooner, M. Seielstad, Y.Y.T., J.C.C., E.S.T. and M.I.M.

MAT2D Consortium samples, genotyping, analysis and management:

J.E.B., G.I.B., J.E., S. Krithika, J. Kumate, A.V.-S., N.J.C., M. Cruz, C.L.H. and E.J.P.

Project management:

D.A., D.W.B., Y.S.C., N.J.C., M. Cruz, C.L.H., J. Kooner, J.-Y.L., M. Seielstad, Y.Y.T., M.B., E.J.P., J.C.C., E.S.T., M.I.M. and A.P.M.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

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

- Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identified additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
- Kong, A. *et al.* Parental origin of sequence variants associated with complex diseases. *Nature* **462**, 868–874 (2009).
- 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).
- Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).
- Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
- Rosenberg, N.A. *et al.* Genome-wide association studies in diverse populations. *Nat. Rev. Genet.* **11**, 356–366 (2010).
- Qi, L. *et al.* Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. *Hum. Mol. Genet.* **19**, 2706–2715 (2010).
- 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).
- Shu, X.O. *et al.* Identification of new genetic risk variants for type 2 diabetes. *PLoS Genet.* **6**, e1001127 (2010).
- 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).
- Cho, Y.S. *et al.* Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. *Nat. Genet.* **44**, 67–72 (2012).
- Li, H. *et al.* A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. *Diabetes* **62**, 291–298 (2013).
- Kooner, J.S. *et al.* Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat. Genet.* **43**, 984–989 (2011).
- Taibassum, R. *et al.* Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes* **62**, 977–986 (2013).
- Parra, E.J. *et al.* Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American samples from Starr County, Texas. *Diabetologia* **54**, 2038–2046 (2011).
- Palmer, N.D. *et al.* A genome-wide association search for type 2 diabetes genes in African Americans. *PLoS ONE* **7**, e29202 (2012).
- Waters, K.M. *et al.* Consistent association of type 2 diabetes risk variants found in Europeans in diverse racial and ethnic groups. *PLoS Genet.* **6**, e1001078 (2010).
- Saxena, R. *et al.* Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. *Am. J. Hum. Genet.* **90**, 410–425 (2012); erratum **90**, 753 (2012).
- Cooper, R.S., Tayo, B. & Zhu, X. Genome-wide association studies: implications for multi-ethnic sciences. *Hum. Mol. Genet.* **17**, R151–R155 (2008).
- Zaitlen, N., Pasaniciuc, B., Gur, T., Ziv, E. & Halperin, E. Leveraging genetic variability across populations for the identification of causal variants. *Am. J. Hum. Genet.* **86**, 23–33 (2010).
- Franceschini, N. *et al.* Discovery and fine-mapping of serum protein loci through transethnic meta-analysis. *Am. J. Hum. Genet.* **91**, 744–753 (2012).
- The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
- The International HapMap Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
- Dickson, S.P., Wang, K., Krantz, I., Hakonarson, H. & Goldstein, D.B. Rare variants create synthetic genome-wide associations. *PLoS Biol.* **8**, e1000294 (2010).
- The 1000 Genomes Project Consortium. A map of human genome variation from population scale sequencing. *Nature* **467**, 1061–1073 (2010).
- Voight, B.F. *et al.* The MetaboChip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet.* **8**, e1002793 (2012).
- Bradfield, J.P. *et al.* A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet.* **7**, e1002293 (2011).
- Cervin, C. *et al.* Genetic similarities between latent autoimmune diabetes in adults, type 1 diabetes, and type 2 diabetes. *Diabetes* **57**, 1433–1437 (2008).
- Grant, S.F., Hakonarson, H. & Schwartz, S. Can the genetics of type 1 and type 2 diabetes shed light on the genetics of latent autoimmune diabetes in adults? *Endocr. Rev.* **31**, 183–193 (2010).
- Scott, R.A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat. Genet.* **44**, 991–1005 (2012).
- Richards, J.B. *et al.* A genome-wide association study reveals variants in ARL15 that influence adiponectin levels. *PLoS Genet.* **5**, e1000768 (2009).
- Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
- Heid, I.M. *et al.* Meta-analysis identifies 12 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42**, 949–960 (2010).
- Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
- Ng, P.C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).

36. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
37. Lee, E.K. *et al.* RNA-binding protein HuD controls insulin translation. *Mol. Cell* **45**, 826–835 (2012).
38. Trynka, G. *et al.* Chromatin marks identify critical cell types for fine-mapping complex trait variants. *Nat. Genet.* **45**, 124–130 (2013).
39. Morris, A.P. Transethnic meta-analysis of genomewide association studies. *Genet. Epidemiol.* **35**, 809–822 (2011).
40. Wang, X. *et al.* Comparing methods for performing trans-ethnic meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **22**, 2303–2311 (2013).
41. Maller, J.B. *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
42. 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 (2012).
43. Fogarty, M.P., Panhuis, T.M., Vadlamudi, S., Buchkovich, M.L. & Mohlke, K.L. Allele-specific transcriptional activity at type 2 diabetes-associated single nucleotide polymorphisms in regions of pancreatic islet open chromatin at the JAZF1 locus. *Diabetes* **62**, 1756–1762 (2013).
44. Nicolson, T.J. *et al.* Insulin storage and glucose homeostasis in mice null for the granule zinc transporter ZnT8 and studies of the type 2 diabetes-associated variants. *Diabetes* **58**, 2070–2083 (2009).
45. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
46. Sung, Y.J., Wang, L., Rankinen, T., Bouchard, C. & Rao, D.C. Performance of genotype imputations using data from the 1000 Genomes Project. *Hum. Hered.* **73**, 18–25 (2012).
47. Zheng, H.F., Ladouceur, M., Greenwood, C.M. & Richards, J.B. Effect of genome-wide genotyping and reference panels on rare variant imputation. *J. Genet. Genomics* **39**, 545–550 (2012).
48. Nelson, S.C. *et al.* Imputation-based genomic coverage assessments of current human genotyping arrays. *G3 (Bethesda)* **3**, 1795–1080 (2013).
49. Stahl, E.A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat. Genet.* **44**, 483–489 (2012).

The authors of this paper are:

Anubha Mahajan^{1,2,20}, Min Jin Go^{2,220}, Weihua Zhang^{3,220}, Jennifer E Below^{4,220}, Kyle J Gaulton^{1,220}, Teresa Ferreira¹, Momoko Horikoshi^{1,5}, Andrew D Johnson⁶, Maggie C Y Ng^{7,8}, Inga Prokopenko^{1,5,9}, Danish Saleheen^{10,11}, Xu Wang¹², Eleftheria Zeggini¹³, Goncalo R Abecasis¹⁴, Linda S Adair¹⁵, Peter Almgren¹⁶, Mustafa Atalay¹⁷, Tin Aung^{18,19}, Damiano Baldassarre^{20,21}, Beverley Balkau^{22,23}, Yuqian Bao²⁴, Anthony H Barnett^{25,26}, Ines Barroso^{13,27,28}, Abdul Basit²⁹, Latonya F Been³⁰, John Beilby^{31–33}, Graeme I Bell^{34,35}, Rafn Benediktsson^{36,37}, Richard N Bergman³⁸, Bernhard O Boehm^{39,40}, Eric Boerwinkle^{41,42}, Lori L Bonnycastle⁴³, Noël Burt⁴⁴, Qiuyin Cai⁴⁵, Harry Campbell^{46,47}, Jason Carey⁴⁴, Stephane Cauchi⁴⁸, Mark Caulfield⁴⁹, Juliana C N Chan⁵⁰, Li-Ching Chang⁵¹, Tien-Jyun Chang⁵², Yi-Cheng Chang⁵², Guillaume Charpentier⁵³, Chien-Hsiun Chen^{51,54}, Han Chen⁵⁵, Yuan-Tsong Chen⁵¹, Kee-Seng Chia^{12,56}, Manickam Chidambaram⁵⁷, Peter S Chines⁴³, Nam H Cho⁵⁸, Young Min Cho⁵⁹, Lee-Ming Chuang^{52,60}, Francis S Collins⁴³, Marilyn C Cornelis⁶¹, David J Couper⁶², Andrew T Crenshaw⁴⁴, Rob M van Dam^{61,63}, John Danesh¹⁰, Debashish Das⁶⁴, Ulf de Faire⁶⁵, George Dedoussis⁶⁶, Panos Deloukas¹³, Antigone S Dimas^{1,67,68}, Christian Dina^{69–71}, Alex S F Doney^{72,73}, Peter J Donnelly^{1,74}, Mozhgan Dorkhan¹⁶, Cornelia van Duijn^{75,76}, José Dupuis^{6,55}, Sarah Edkins¹³, Paul Elliott^{3,77}, Valur Emilsson⁷⁸, Raimund Erbel⁷⁹, Johan G Eriksson^{80–83}, Jorge Escobedo⁸⁴, Tonu Esko^{85–88}, Elodie Eury⁴⁸, Jose C Florez^{87–91}, Pierre Fontanillas⁴⁴, Nita G Forouhi⁹², Tom Forsen^{81,93}, Caroline Fox^{6,94}, Ross M Fraser⁴⁶, Timothy M Frayling⁹⁵, Philippe Froguel^{9,48}, Philippe Frossard¹¹, Yutang Gao⁹⁶, Karl Gertow^{97,98}, Christian Gieger⁹⁹, Bruna Gigante⁶⁵, Harald Grallert^{100–103}, George B Grant⁴⁴, Leif C Groop¹⁶, Christopher J Groves⁵, Elin Grundberg¹⁰⁴, Candace Guiducci⁴⁴, Anders Hamsten^{97,98}, Bok-Ghee Han², Kazuo Hara¹⁰⁵, Neelam Hassanali⁵, Andrew T Hattersley¹⁰⁶, Caroline Hayward⁴⁷, Asa K Hedman¹, Christian Herder¹⁰⁷, Albert Hofman⁷⁵, Oddgeir L Holmen¹⁰⁸, Kees Hovingh¹⁰⁹, Astradur B Hreidarsson³⁷, Cheng Hu²⁴, Frank B Hu^{61,110}, Jennie Hui^{31,111}, Steve E Humphries¹¹², Sarah E Hunt¹³, David J Hunter^{61,110,113}, Kristian Hveem¹⁰⁸, Zafar I Hydrie²⁹, Hiroshi Ikegami¹¹⁴, Thomas Illig^{100,115}, Erik Ingelsson^{1,116}, Muhammed Islam¹¹⁷, Bo Isomaa^{83,118}, Anne U Jackson¹⁴, Tazeen Jafar^{117,119}, Alan James^{31,120,121}, Weiping Jia²⁴, Karl-Heinz Jöckel¹²², Anna Jonsson¹⁶, Jeremy B M Jowett¹²³, Takashi Kadowaki¹⁰⁵, Hyun Min Kang¹⁴, Stavroula Kanoni¹³, Wen Hong L Kao¹²⁴, Sekar Kathiresan^{44,90,125}, Norihiro Kato¹²⁶, Prasad Katulanda^{5,127}, Sirkka M Keinänen-Kiukaanniemi^{128,129}, Ann M Kelly^{25,26}, Hassan Khan¹⁰, Kay-Tee Khaw¹⁰, Chiea-Chuen Khor^{18,56,130}, Hyung-Lae Kim¹³¹, Sangsoo Kim¹³², Young Jin Kim², Leena Kinnunen¹³³, Norman Klopp^{100,115}, Augustine Kong¹³⁴, Eeva Korpi-Hyövälti¹³⁵, Sudhir Kowlessur¹³⁶, Peter Kraft^{61,113}, Jasmina Kravic¹⁶, Malene M Kristensen¹²³, S Krithika¹³⁷, Ashish Kumar¹, Jesus Kumate¹³⁸, Johanna Kuusisto¹³⁹, Soo Heon Kwak⁵⁹, Markku Laakso¹³⁹, Vasiliki Lagou¹, Timo A Lakka^{17,140}, Claudia Langenberg⁹², Cordelia Langford¹³, Robert Lawrence¹⁴¹, Karin Leander⁶⁵, Jen-Mai Lee¹², Nanette R Lee¹⁴², Man Li¹²⁴, Xinzhong Li¹⁴³, Yun Li^{144,145}, Junbin Liang¹⁴⁶, Samuel Liju⁵⁷, Wei-Yen Lim⁵⁶, Lars Lind¹⁴⁷, Cecilia M Lindgren^{1,44}, Eero Lindholm¹⁶, Ching-Ti Liu⁵⁵, Jian Jun Liu¹³⁰, Stéphane Lobbens⁴⁸, Jirong Long⁴⁵, Ruth J F Loos^{92,148–150}, Wei Lu¹⁵¹, Jian'an Luan⁹², Valeriya Lyssenko¹⁶, Ronald C W Ma⁵⁰, Shiro Maeda⁶⁰, Reedik Mägi⁸⁵, Satu Männistö⁸⁰, David R Matthews⁵, James B Meigs^{89,152}, Olle Melander¹⁶, Andres Metspalu^{85,86}, Julia Meyer⁹⁹, Ghazala Mirza¹, Evelin Mihailov⁸⁵, Susanne Moebus¹²², Viswanathan Mohan^{57,153}, Karen L Mohlke¹⁴⁴, Andrew D Morris^{72,73},

Thomas W Mühleisen^{154,155}, Martina Müller-Nurasyid^{99,156,157}, Bill Musk^{31,111,121,158}, Jiro Nakamura¹⁵⁹, Eitaro Nakashima^{159,160}, Pau Navarro⁴⁷, Peng-Keat Ng¹², Alexandra C Nica⁶⁷, Peter M Nilsson¹⁶, Inger Njølstad¹⁶¹, Markus M Nöthen^{154,155}, Keizo Ohnaka¹⁶², Twee Hee Ong¹², Katharine R Owen^{5,163}, Colin N A Palmer^{72,73}, James S Pankow¹⁶⁴, Kyong Soo Park^{59,165}, Melissa Parkin⁴⁴, Sonali Pechlivanis¹²², Nancy L Pedersen¹⁶⁶, Leena Peltonen^{13,44,80,167,219}, John R B Perry^{1,95}, Annette Peters¹⁶⁸, Janani M Pinidiyapathirage¹⁶⁹, Carl G P Platou^{108,170}, Simon Potter¹³, Jackie F Price⁴⁶, Lu Qi^{61,110}, Venkatesan Radha⁵⁷, Loukianos Rallidis¹⁷¹, Asif Rasheed¹¹, Wolfgang Rathmann¹⁷², Rainer Rauramaa^{140,173}, Soumya Raychaudhuri^{44,174,175}, N William Rayner^{1,5}, Simon D Rees^{25,26}, Emil Rehnberg¹⁶⁶, Samuli Ripatti^{13,80,167}, Neil Robertson^{1,5}, Michael Roden^{107,176,177}, Elizabeth J Rossin^{44,87,178,179}, Igor Rudan⁴⁶, Denis Rybin¹⁸⁰, Timo E Saaristo^{181,182}, Veikko Salomaa⁸⁰, Juha Saltevo¹⁸³, Maria Samuel¹¹, Dharambir K Sanghera³⁰, Jouko Saramies¹⁸⁴, James Scott¹⁴³, Laura J Scott¹⁴, Robert A Scott⁹², Ayellet V Segre^{44,89,90}, Joban Sehmi^{64,143}, Bengt Sennblad^{97,98}, Nabi Shah¹¹, Sonia Shah¹⁸⁵, A Samad Shera¹⁸⁶, Xiao Ou Shu⁴⁵, Alan R Shuldiner¹⁸⁷⁻¹⁸⁹, Gunnar Sigurðsson^{37,78}, Eric Sijbrands¹⁹⁰, Angela Silveira^{97,98}, Xueling Sim^{14,56}, Suthesh Sivapalaratnam¹⁰⁹, Kerrin S Small^{13,104}, Wing Yee So⁵⁰, Alena Stančáková¹³⁹, Kari Stefansson^{36,134}, Gerald Steinbach¹⁹¹, Valgerdur Steinthorsdottir¹³⁴, Kathleen Stirrups¹³, Rona J Strawbridge^{97,98}, Heather M Stringham¹⁴, Qi Sun^{61,110}, Chen Suo⁵⁶, Ann-Christine Syvänen¹⁹², Ryoichi Takayanagi¹⁹³, Fumihiko Takeuchi¹²⁶, Wan Ting Tay¹⁸, Tanya M Teslovich¹⁴, Barbara Thorand¹⁶⁸, Gudmar Thorleifsson¹³⁴, Unnur Thorsteinsdottir^{36,134}, Emmi Tikkanen^{80,167}, Joseph Trakalo¹, Elena Tremoli^{20,21}, Mieke D Trip¹⁰⁹, Fuu Jen Tsai⁵⁴, Tiinamaija Tuomi^{83,194}, Jaakko Tuomilehto^{133,195-197}, Andre G Uitterlinden^{75,76,190}, Adan Valladares-Salgado¹⁹⁸, Sailaja Vedantam^{87,88}, Fabrizio Veglia²⁰, Benjamin F Voight^{44,199}, Congrong Wang²⁴, Nicholas J Wareham⁹², Roman Wennauer¹⁹⁰, Ananda R Wickremasinghe¹⁶⁹, Tom Wilsgaard¹⁶¹, James F Wilson^{46,47}, Steven Wiltshire^{1,219}, Wendy Winckler⁴⁴, Tien Yin Wong^{18,19,200}, Andrew R Wood⁹⁵, Jer-Yuarn Wu^{51,54}, Ying Wu¹⁴⁴, Ken Yamamoto²⁰¹, Toshimasa Yamauchi¹⁰⁵, Mingyu Yang¹⁴⁶, Loic Yengo^{48,202}, Mitsuhiro Yokota²⁰³, Robin Young¹⁰, Delilah Zabaneh¹⁸⁵, Fan Zhang¹⁴⁶, Rong Zhang²⁴, Wei Zheng⁴⁵, Paul Z Zimmet¹²³, David Altshuler^{44,89,90,204-206,221}, Donald W Bowden^{7,8,207,208,221}, Yoon Shin Cho^{209,221}, Nancy J Cox^{34,35,221}, Miguel Cruz^{198,221}, Craig L Hanis^{210,221}, Jaspal Kooner^{64,143,211,221}, Jong-Young Lee^{2,221}, Mark Seielstad^{130,212,213,221}, Yik Ying Teo^{12,56,130,214,215,221}, Michael Boehnke^{14,221}, Esteban J Parra^{137,221}, John C Chambers^{3,64,211,221}, E Shyong Tai^{12,216,217,221}, Mark I McCarthy^{1,5,163,221} & Andrew P Morris^{1,218,221}

¹Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ²Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Cheongwon-gun, Gangoe-myeon, Yeonje-ri, Korea. ³Epidemiology and Biostatistics, Imperial College London, London, UK. ⁴Department of Genome Sciences, University of Washington, Seattle, Washington, USA. ⁵Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, UK. ⁶National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. ⁷Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. ⁸Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. ⁹Genomics of Common Disease, Imperial College London, Hammersmith Hospital, London, UK. ¹⁰Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. ¹¹Center for Non-Communicable Diseases Pakistan, Karachi, Pakistan. ¹²Department of Epidemiology and Public Health, National University of Singapore, Singapore. ¹³Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. ¹⁴Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA. ¹⁵Department of Nutrition, University of North Carolina, Chapel Hill, North Carolina, USA. ¹⁶Lund University Diabetes Centre, Department of Clinical Science Malmö, Scania University Hospital, Lund University, Malmö, Sweden. ¹⁷Institute of Biomedicine, University of Eastern Finland, Kuopio Campus, Kuopio, Finland. ¹⁸Singapore Eye Research Institute, Singapore National Eye Centre, Singapore. ¹⁹Department of Ophthalmology, National University of Singapore, Singapore. ²⁰Centro Cardiologico Monzino, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Milan, Italy. ²¹Department of Pharmaceutical Sciences, University of Milano, Milan, Italy. ²²Institut National de la Santé et de la Recherche Médicale (INSERM) Centre de recherche en Epidémiologie et Santé des Populations (CESP) U1018, Villejuif, France. ²³University Paris Sud 11, UMRS 1018, Villejuif, France. ²⁴Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China. ²⁵College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK. ²⁶BioMedical Research Centre, Heart of England National Health Service (NHS) Foundation Trust, Birmingham, UK. ²⁷University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. ²⁸National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. ²⁹Baqai Institute of Diabetology and Endocrinology, Karachi, Pakistan. ³⁰Department of Pediatrics, Section of Genetics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA. ³¹Busselton Population Medical Research Institute, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. ³²PathWest Laboratory Medicine of Western Australia, Queen Elizabeth II Medical Centre, Nedlands, Western Australia, Australia. ³³School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, Western Australia, Australia. ³⁴Department of Medicine, University of Chicago, Chicago, Illinois, USA. ³⁵Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. ³⁶Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ³⁷Landspítali University Hospital, Reykjavik, Iceland. ³⁸Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. ³⁹Department of Internal Medicine, Division of Endocrinology and Diabetes, University Medical Centre Ulm, Ulm, Germany. ⁴⁰Lee Kong Chian (LKC) School of Medicine, Nanyang Technological University, Singapore and Imperial College London, London, UK. ⁴¹Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA. ⁴²Human Genome Sequencing Center at Baylor College of Medicine, Houston, Texas, USA. ⁴³National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA. ⁴⁴Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA. ⁴⁵Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. ⁴⁶Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, UK. ⁴⁷Medical Research Council (MRC) Institute of Genetics and Molecular Medicine at the University of Edinburgh, Western General Hospital, Edinburgh, UK. ⁴⁸Centre National de la Recherche Scientifique (CNRS)-Unité Mixte de Recherche (UMR)-8199, Institute of Biology and Lille 2 University, Pasteur Institute, Lille, France.

⁴⁹Clinical Pharmacology and Barts and the London Genome Centre, William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London, UK. ⁵⁰Department of Medicine and Therapeutics, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China. ⁵¹Institute of Biomedical Sciences, Academia Sinica, Nankang, Taipei, Taiwan. ⁵²Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan. ⁵³Endocrinology-Diabetology Unit, Corbeil-Essonnes Hospital, Corbeil-Essonnes, France. ⁵⁴School of Chinese Medicine, China Medical University, Taichung, Taiwan. ⁵⁵Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. ⁵⁶Centre for Molecular Epidemiology, National University of Singapore, Singapore. ⁵⁷Department of Molecular Genetics, Madras Diabetes Research Foundation–Indian Council of Medical Research (ICMR) Advanced Centre for Genomics of Diabetes, Chennai, India. ⁵⁸Department of Preventive Medicine, Ajou University School of Medicine, Suwon, Korea. ⁵⁹Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea. ⁶⁰Graduate Institute of Clinical Medicine, National Taiwan University School of Medicine, Taipei, Taiwan. ⁶¹Department of Nutrition and Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. ⁶²Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. ⁶³Saw Swee Hock School of Public Health, National University of Singapore, Singapore. ⁶⁴Ealing Hospital NHS Trust, Middlesex, UK. ⁶⁵Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. ⁶⁶Department of Dietetics-Nutrition, Harokopio University, Athens, Greece. ⁶⁷Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. ⁶⁸Biomedical Sciences Research Center 'Alexander Fleming', Vari, Greece. ⁶⁹INSERM UMR 1087, Nantes, France. ⁷⁰CNRS UMR 6291, Nantes, France. ⁷¹Nantes University, Nantes, France. ⁷²Diabetes Research Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee, UK. ⁷³Pharmacogenomics Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee, UK. ⁷⁴Department of Statistics, University of Oxford, Oxford, UK. ⁷⁵Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands. ⁷⁶Netherland Genomics Initiative, Netherlands Consortium for Healthy Ageing and Centre for Medical Systems Biology, Rotterdam, The Netherlands. ⁷⁷MRC–Health Protection Agency (HPA) Centre for Environment and Health, Imperial College London, London, UK. ⁷⁸Icelandic Heart Association, Kopavogur, Iceland. ⁷⁹Clinic of Cardiology, West German Heart Centre, University Hospital of Essen, University Duisburg-Essen, Essen, Germany. ⁸⁰Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. ⁸¹Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland. ⁸²Unit of General Practice, Helsinki University General Hospital, Helsinki, Finland. ⁸³Folkhälsan Research Center, Helsinki, Finland. ⁸⁴Unidad de Investigación en Epidemiología Clínica, Hospital General Regional I, Dr. Carlos MacGregor, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico. ⁸⁵Estonian Genome Center, University of Tartu, Tartu, Estonia. ⁸⁶Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. ⁸⁷Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. ⁸⁸Division of Genetics and Endocrinology, Children's Hospital, Boston, Massachusetts, USA. ⁸⁹Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. ⁹⁰Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. ⁹¹Diabetes Research Center, Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. ⁹²MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. ⁹³Vaasa Health Care Centre, Vaasa, Finland. ⁹⁴Division of Endocrinology and Metabolism, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. ⁹⁵Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter, UK. ⁹⁶Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China. ⁹⁷Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden. ⁹⁸Center for Molecular Medicine, Karolinska University Hospital Solna, Stockholm, Sweden. ⁹⁹Institute of Genetic Epidemiology, Helmholtz Zentrum Muenchen, Neuherberg, Germany. ¹⁰⁰Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, Neuherberg, Germany. ¹⁰¹Clinical Cooperation Group Diabetes, Ludwig-Maximilians-Universität Muenchen and Helmholtz Zentrum Muenchen, Muenchen, Germany. ¹⁰²Clinical Cooperation Group Nutrigenomics and Type 2 Diabetes, Technical University Muenchen and Helmholtz Zentrum Muenchen, Muenchen, Germany. ¹⁰³German Center for Diabetes Research (DZD), Neuherberg, Germany. ¹⁰⁴Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. ¹⁰⁵Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. ¹⁰⁶Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter, UK. ¹⁰⁷Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ¹⁰⁸HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway. ¹⁰⁹Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. ¹¹⁰Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. ¹¹¹School of Population Health, The University of Western Australia, Nedlands, Western Australia, Australia. ¹¹²Institute of Cardiovascular Science, University College London, London, UK. ¹¹³Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. ¹¹⁴Department of Endocrinology, Metabolism and Diabetes, Kinki University School of Medicine, Osaka-sayama, Osaka, Japan. ¹¹⁵Hannover Unified Biobank, Hannover Medical School, Hannover, Germany. ¹¹⁶Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University Hospital, Uppsala, Sweden. ¹¹⁷Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan. ¹¹⁸Department of Social Services and Health Care, Jakobstad, Finland. ¹¹⁹Department of Medicine, Aga Khan University, Karachi, Pakistan. ¹²⁰Department of Pulmonary Physiology and Sleep Medicine, West Australian Sleep Disorders Research Institute, Queen Elizabeth II Medical Centre, Hospital Avenue, Nedlands, Western Australia, Australia. ¹²¹School of Medicine and Pharmacology, University of Western Australia, Nedlands, Western Australia, Australia. ¹²²Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, Germany. ¹²³Baker International Diabetes Institute (IDI) Heart and Diabetes Institute, Melbourne, Victoria, Australia. ¹²⁴Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. ¹²⁵Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA. ¹²⁶Research Institute, National Center for Global Health and Medicine, Shinjuku-ku, Tokyo, Japan. ¹²⁷Diabetes Research Unit, Department of Clinical Medicine, University of Colombo, Colombo, Sri Lanka. ¹²⁸Faculty of Medicine, Institute of Health Sciences, University of Oulu, Oulu, Finland. ¹²⁹Unit of General Practice, Oulu University Hospital, Oulu, Finland. ¹³⁰Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore. ¹³¹Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Korea. ¹³²School of Systems Biomedical Science, Soongsil University, Dongjak-gu, Seoul, Korea. ¹³³Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland. ¹³⁴deCODE Genetics, Reykjavik, Iceland. ¹³⁵South Ostrobothnia Central Hospital, Seinäjoki, Finland. ¹³⁶Ministry of Health, Port Louis, Mauritius. ¹³⁷Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada. ¹³⁸Fundacion IMSS, Mz. Poniente, Col. Juarez, Deleg. Cuauhtemoc, Mexico City, Mexico. ¹³⁹Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. ¹⁴⁰Kuopio Research Institute of Exercise Medicine, Kuopio, Finland. ¹⁴¹Centre for Genetic Epidemiology and Biostatistics, The University of Western Australia, Nedlands, Western Australia, Australia. ¹⁴²Office of Population Studies Foundation Inc., University of San Carlos, Cebu City, Philippines. ¹⁴³National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK. ¹⁴⁴Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA. ¹⁴⁵Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, USA. ¹⁴⁶Beijing Genomics Institute, Shenzhen, China. ¹⁴⁷Department of Medical Sciences, Uppsala University, Akademiska Sjukhuset, Uppsala, Sweden. ¹⁴⁸Charles R. Bronfman Institute for Personalized Medicine, Mount Sinai School of Medicine, New York, New York, USA. ¹⁴⁹Child Health and Development Institute, Mount Sinai School of Medicine, New York, New York, USA. ¹⁵⁰Department of Preventive Medicine, Mount Sinai School of Medicine, New York, New York, USA. ¹⁵¹Shanghai Institute of Preventive Medicine, Shanghai, China. ¹⁵²General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA. ¹⁵³Dr Mohan's Diabetes Specialties Centre, Chennai, India. ¹⁵⁴Institute of Human Genetics, University of Bonn, Bonn, Germany. ¹⁵⁵Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany. ¹⁵⁶Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. ¹⁵⁷Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-University, Munich, Germany. ¹⁵⁸Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. ¹⁵⁹Department of Internal Medicine, Division of Endocrinology and Diabetes, Nagoya University Graduate School of Medicine, Nagoya, Japan. ¹⁶⁰Department of Diabetes and Endocrinology, Chubu Rosai Hospital, Nagoya, Japan. ¹⁶¹Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. ¹⁶²Department of Geriatric Medicine, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan. ¹⁶³Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Oxford, UK. ¹⁶⁴Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, USA. ¹⁶⁵World 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. ¹⁶⁶Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ¹⁶⁷Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland. ¹⁶⁸Institute of Epidemiology II, Helmholtz Zentrum Muenchen, Neuherberg, Germany. ¹⁶⁹Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka. ¹⁷⁰Department of Internal Medicine, Levanger Hospital, Nord-Trøndelag Health Trust, Levanger, Norway. ¹⁷¹University General Hospital Attikon, 1 Rimini, Athens, Greece. ¹⁷²Institute of Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

¹⁷³Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. ¹⁷⁴Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. ¹⁷⁵Partners Center for Personalized Genomic Medicine, Boston, Massachusetts, USA. ¹⁷⁶Department of Endocrinology and Diabetology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ¹⁷⁷Department of Metabolic Diseases, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ¹⁷⁸Health Science and Technology MD Program, Harvard University and Massachusetts Institute of Technology, Boston, Massachusetts, USA. ¹⁷⁹Harvard Biological and Biomedical Sciences Program, Harvard University, Boston, Massachusetts, USA. ¹⁸⁰Boston University Data Coordinating Center, Boston, Massachusetts, USA. ¹⁸¹Finnish Diabetes Association, Tampere, Finland. ¹⁸²Pirkanmaa Hospital District, Tampere, Finland. ¹⁸³Department of Medicine, Central Finland Central Hospital, Jyväskylä, Finland. ¹⁸⁴South Karelia Central Hospital, Lappeenranta, Finland. ¹⁸⁵University College London Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK. ¹⁸⁶Diabetic Association Pakistan, Karachi, Pakistan. ¹⁸⁷Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA. ¹⁸⁸Geriatric Research Education and Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, Maryland, USA. ¹⁸⁹Program in Personalised and Genomic Medicine, University of Maryland School of Medicine, Baltimore, Maryland, USA. ¹⁹⁰Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands. ¹⁹¹Department of Clinical Chemistry and Central Laboratory, University of Ulm, Ulm, Germany. ¹⁹²Molecular Medicine, Department of Medical Sciences, Uppsala University, Uppsala, Sweden. ¹⁹³Department of Internal Medicine and Bioregulatory Science, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan. ¹⁹⁴Department of Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. ¹⁹⁵Instituto de Investigacion Sanitaria del Hospital Universitario LaPaz (IdiPAZ), Madrid, Spain. ¹⁹⁶Centre for Vascular Prevention, Danube-University Krems, Krems, Austria. ¹⁹⁷Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia. ¹⁹⁸Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, IMSS, Av. Cuauhtemoc 330, Col. Doctores, Mexico City, Mexico. ¹⁹⁹University of Pennsylvania–Perelman School of Medicine, Department of Pharmacology, Philadelphia, Pennsylvania, USA. ²⁰⁰Centre for Eye Research Australia, University of Melbourne, East Melbourne, Victoria, Australia. ²⁰¹Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka, Japan. ²⁰²University Lille 1, Laboratory of Mathematics, CNRS-UMR 8524, MODAL team, INRIA Lille Nord-Europe, Villeneuve-d'Ascq, France. ²⁰³Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya, Japan. ²⁰⁴Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. ²⁰⁵Department of Molecular Biology, Harvard Medical School, Boston, Massachusetts, USA. ²⁰⁶Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. ²⁰⁷Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. ²⁰⁸Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. ²⁰⁹Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, Korea. ²¹⁰Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA. ²¹¹Imperial College Healthcare NHS Trust, London, UK. ²¹²Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA. ²¹³Blood Systems Research Institute, San Francisco, California, USA. ²¹⁴Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore. ²¹⁵Department of Statistics and Applied Probability, National University of Singapore, Singapore. ²¹⁶Department of Medicine, National University of Singapore, Singapore. ²¹⁷Duke–National University of Singapore Graduate Medical School, Singapore. ²¹⁸Department of Biostatistics, University of Liverpool, Liverpool, UK. ²¹⁹Deceased. ²²⁰These authors contributed equally to this work. ²²¹These authors jointly directed this work. Correspondence should be addressed to M.B. (boehnke@umich.edu), E.J.P. (esteban.parra@utoronto.ca), J.C.C. (john.chambers@imperial.ac.uk), E.S.T. (e_shyong_tai@nuhs.edu.sg), M.I.M. (mark.mccarthy@drli.ox.ac.uk) or A.P.M. (a.p.morris@liverpool.ac.uk).

ONLINE METHODS

Ancestry-specific GWAS meta-analyses. Ancestry-specific meta-analyses have been previously performed by the DIAGRAM Consortium (12,171 cases and 56,862 controls; European ancestry)⁵; the AGEN-T2D Consortium (6,952 cases and 11,865 controls; east Asian ancestry)¹¹; the SAT2D Consortium (5,561 cases and 14,458 controls; south Asian ancestry)¹³; and the MAT2D Consortium (1,804 cases and 779 controls; Mexican and Mexican American ancestry)¹⁵. Further details of the samples and methods employed within each ancestry group are presented in the corresponding consortium papers^{5,11,13,15}. Briefly, individuals were assayed with a range of genotyping products, with sample and SNP quality control (QC) undertaken within each study (**Supplementary Tables 1 and 2**). Each GWAS scaffold was imputed up to 2.5 million autosomal SNPs using reference panels from Phase II/III HapMap^{22,23} (**Supplementary Table 2**). Each SNP with MAF > 1%, (or MAF > 5% in the Mexican and Mexican American ancestry GWAS because of smaller sample size) and passing QC was tested for association with T2D under an additive model after adjustment for study-specific covariates (**Supplementary Table 2**). Covariate adjustments were not harmonized within or between consortia because of variation in study design and recorded non-genetic risk factors. The results of each GWAS were corrected for population structure with genomic control⁵⁰ (unless $\lambda_{GC} < 1$). Association summary statistics from GWAS within each ancestry group were then combined by fixed-effects meta-analysis. The results of each ethnic-specific meta-analysis were then corrected by a second round of genomic control: European ancestry ($\lambda_{GC} = 1.10$); east Asian ancestry ($\lambda_{GC} = 1.05$); south Asian ancestry ($\lambda_{GC} = 1.02$); Mexican and Mexican American ancestry ($\lambda_{GC} = 1.01$).

Trans-ethnic discovery-stage GWAS meta-analysis. Association summary statistics from each ancestry-specific meta-analysis were combined in a fixed-effects inverse-variance-weighted meta-analysis (in a total of 26,488 cases and 83,964 controls). The association results of the trans-ethnic meta-analysis were corrected by genomic control⁵⁰ ($\lambda_{GC} = 1.05$).

Heterogeneity analyses. For each previously reported lead SNP at an established T2D susceptibility locus, we assessed heterogeneity in allelic effects between the ethnic-specific meta-analyses by means of Cochran's Q statistic⁵¹ (**Supplementary Table 3**). Among the 52 SNPs passing QC in all four ethnic-specific meta-analyses, we identified those that showed the same direction of effect across all ancestry groups and evaluated the significance of the excess in concordance (12.5% expected) with a one-sided binomial test.

Concordance analyses. We identified SNPs passing QC and with MAF > 1% in all four ethnic-specific meta-analyses. We excluded variants in the 69 established autosomal T2D susceptibility loci, defined as 500 kb upstream and 500 kb downstream of the previously reported lead SNPs. We also excluded AT/GC SNPs to eliminate bias due to strand misalignment between ethnic-specific meta-analyses. Among the remaining SNPs, we selected an independent subset with nominal evidence of association ($P \leq 0.001$) with T2D from the European ancestry meta-analysis and separated by at least 500 kb. For each independent SNP, we identified the T2D risk allele from the European ancestry meta-analysis and determined the direction of effect in the east Asian, south Asian and Mexican and Mexican American ancestry meta-analyses. We calculated the proportion of these SNPs that had the same direction of effect for the European ancestry risk allele and the significance of the excess in concordance (50% expected) with a one-sided binomial test. We repeated this analysis for SNPs with weaker evidence of association with T2D from the European ancestry meta-analysis: $0.001 < P \leq 0.01$; $0.01 < P \leq 0.5$; and $0.5 < P \leq 1$ (**Table 1**). We then repeated these analyses using the east Asian, south Asian and Mexican and Mexican American ancestry meta-analyses, in turn, to identify subsets of independent T2D risk alleles and assessed concordance into the other ethnic groups (**Supplementary Table 4**).

European ancestry validation-stage meta-analysis. The previously published validation meta-analysis consisted of 21,491 cases and 55,647 controls of European ancestry from the DIAGRAM Consortium⁵, all genotyped with the Metachip²⁶ (**Supplementary Table 1**). We excluded the Pakistan Risk Of Myocardial Infarction Study (PROMIS) from the validation meta-analysis

to avoid overlap with a subset of the same individuals contributing to the SAT2D Consortium meta-analysis¹³. Full details of the samples and methods employed in the validation meta-analysis are presented in the DIAGRAM Consortium paper⁵. Briefly, sample and SNP QC were undertaken within each study (**Supplementary Table 2**). Each high-quality SNP (MAF > 1%) was tested for association with T2D under an additive model after adjustment for study-specific covariates (**Supplementary Table 2**). Association summary statistics for each study were corrected using the genomic control inflation factor obtained from a subset of 3,598 'QT interval' replication SNPs^{5,26} (unless $\lambda_{QT} < 1$). These statistics were then combined in a fixed-effects inverse-variance-weighted meta-analysis and were corrected by a second round of genomic control ($\lambda_{QT} = 1.19$).

Combined meta-analysis. We selected lead SNPs at 33 new loci with suggestive evidence of association ($P < 10^{-5}$) from the trans-ethnic discovery GWAS meta-analysis for *in silico* follow up in the European ancestry validation meta-analysis. Of these SNPs, 16 were genotyped directly on Metachip, and 10 more had a proxy (CEU and CHB+JPT HapMap $r^2 \geq 0.2$). For these 26 SNPs, association summary statistics from the discovery and validation meta-analyses were combined in a fixed-effects inverse-variance-weighted meta-analysis (**Supplementary Table 5**). The combined meta-analysis consisted of 47,979 T2D cases and 139,611 controls. Heterogeneity in allelic effects between the two stages of the combined meta-analysis was assessed by means of Cochran's Q statistic⁵¹.

Sensitivity to covariate adjustment. We identified 19 studies (11,327 cases and 31,342 controls) from the European ancestry validation meta-analysis that adjusted for only age, sex (unless male or female specific) and population structure where necessary (**Supplementary Table 2**): AMC-PAS; BHS; DILGOM; EAS; EGCUT; EMIL-ULM; EPIC; FUSION Stage 2; D2D2007; Dr's Extra; HUNT; METSIM (male specific); HNR; IMPROVE; KORAGEN Stage 2; PIVUS; THISEAS; ULSAM (male specific); and WARREN2. Association summary statistics from each of these studies were then combined in a fixed-effects inverse-variance-weighted meta-analysis, the results of which were subsequently corrected for genomic control ($\lambda_{QT} = 1.12$). The remaining six studies (10,164 cases and 24,305 controls) did not adjust for age and/or sex or included additional covariates to account for BMI or cardiovascular-related disease status (**Supplementary Table 2**): deCODE Stage 2; DUNDEE; GMetS; PMB; SCARFSHEEP; and STR. Association summary statistics from each of these studies were then combined in a fixed-effects inverse-variance-weighted meta-analysis but did not require subsequent correction for genomic control ($\lambda_{QT} = 1.00$). We then tested for heterogeneity in allelic effects between these two sets of studies by means of Cochran's Q statistic⁵¹ (**Supplementary Table 7**).

Association of lead T1D SNPs with T2D. We obtained association summary statistics with T2D from the trans-ethnic meta-analysis for previously reported lead SNPs in established T1D susceptibility loci²⁷ (**Supplementary Table 8**). For each SNP, we aligned the allelic effect on T2D according to the risk allele for T1D (where reported). We also obtained association summary statistics for tags for T1D HLA risk alleles: *HLA-DR3* (rs660895) and *HLA-DR3* (rs2187668).

Association of lead T2D SNPs with metabolic traits. We obtained association summary statistics (P values, directed Z scores and/or allelic effects and corresponding standard errors) for lead SNPs at new T2D susceptibility loci in published European ancestry GWAS meta-analyses of metabolic phenotypes: glycaemic traits^{3,30}, anthropometric measures^{32,33} and plasma lipid concentrations³⁴. We considered glycaemic traits in non-diabetic individuals from the MAGIC Investigators (**Supplementary Table 9**). For FG and FI concentrations (with and without adjustment for BMI), the meta-analysis consisted of up to 133,010 and 108,557 individuals, respectively. For HOMA-B and HOMA-IR, the meta-analysis consisted of up to 37,037 individuals. We considered anthropometric measures from the GIANT Consortium (**Supplementary Table 10**). For BMI and waist-hip ratio adjusted for BMI, the meta-analysis consisted of 123,865 and 77,167 individuals, respectively. We then considered plasma lipid concentrations from the Global Lipids Genetics Consortium (**Supplementary Table 11**). For total cholesterol, high-density lipoprotein cholesterol, low-density

lipoprotein cholesterol and triglycerides, the meta-analysis consisted of up to 100,184 individuals.

Expression analyses. We interrogated public databases and unpublished resources for *cis*-eQTL expression with lead SNPs in the new susceptibility loci in multiple tissues. Details of these resources are summarized in the **Supplementary Note**. The collated results from these resources met study-specific criteria for statistical significance for association with expression. For each transcript associated with the lead T2D SNP (**Supplementary Table 12**), we identified the *cis*-eQTL SNP with the strongest association with expression in the same tissue and subsequently estimated the LD between them using pilot data from the 1000 Genomes Project²⁵ (CEU and CHB+JPT) to assess the coincidence of the signals.

Functional annotation. We identified variants in pilot data from the 1000 Genomes Project²⁵ that are in strong LD (CEU and CHB+JPT $r^2 > 0.8$) with the lead SNPs in the new susceptibility loci for functional annotation. Identified nonsynonymous variants were interrogated for likely downstream functional consequences using SIFT³⁵ (**Supplementary Table 13**). Variants were also assessed for overlap with regions of predicted regulatory function generated by the ENCODE Project³⁶ including: ChromHMM regulatory state definitions from 9 cell lines (GM12878, HepG2, HUVEC, HMEC, HSMM, K562, NHLF, NHEK and hESC); transcription factor binding ChIP sites from 95 cell types; open chromatin (DNaseI hypersensitivity) sites from 125 cell types; transcripts correlated with open chromatin site activity; and sequence motifs from JASPAR, TRANSFAC and *de novo* prediction (**Supplementary Fig. 2**).

Fine-mapping analyses. We used MANTRA³⁹ to fine-map T2D susceptibility loci on the basis of association summary statistics from the meta-analysis of European ancestry GWAS only⁵ and the trans-ethnic meta-analysis of European, east Asian, south Asian and Mexican and Mexican American

ancestry GWAS^{5,11,13,15}. MANTRA allows for trans-ethnic heterogeneity in allelic effects arising as a result of differences in the structure of LD with the causal variant in diverse populations by assigning ancestry groups to 'clusters' according to a Bayesian partition model of relatedness between them, as defined by pairwise genome-wide mean allele frequency differences (**Supplementary Fig. 4**). Evidence in favor of association of each SNP with T2D is measured by a Bayes' factor (BF). We assume a single causal variant for T2D at each locus (defined by the region 500 kb upstream and 500 kb downstream of the lead SNP from the trans-ethnic meta-analysis). We then calculated the posterior probability that the j th SNP is causal among those reported in the meta-analysis:

$$\varphi_j = \frac{BF_j}{\sum_k BF_k}$$

In this expression, BF_j denotes the BF in favor of association of the j th SNP, and the summation in the denominator is over all variants passing QC across the locus⁴¹. A 99% credible set of variants was then constructed by ranking all SNPs according to their BF and combining ranked SNPs until their cumulative posterior probability exceeds 0.99.

SNPs in the 99% credible sets were assessed for enrichment in ChromHMM regulatory state (enhancer, promoter and insulator), DNaseI hypersensitive and transcription factor binding sites using data from the ENCODE Project³⁶. We performed 1,000 permutations by shifting the location of the annotation sites a random distance within 100 kb and recalculated the overlap to obtain empirical P values for enrichment in each annotation category.

50. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).

51. Ioannidis, J.P. *et al.* Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS ONE* **2**, e841 (2007).