Title

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

Journal

Nature Genetics, 46(3)

ISSN

1061-4036

Authors

Mahajan, Anubha
Go, Min Jin
Zhang, Weihua
et al.

Publication Date

2014-03-01

DOI

10.1038/ng.2897

Peer reviewed

HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2014 September 01.

Published in final edited form as:

Nat Genet. 2014 March ; 46(3): 234–244. doi:10.1038/ng.2897.

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, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, Mexican American Type 2 Diabetes (MAT2D) Consortium, Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium[2]

Abstract

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@drl.ox.ac.uk), A.P.M. (a.p.morris@liverpool.ac.uk).

[2]A list of members and affiliations of each consortium appears in the Supplementary Note.

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., 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., 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. Sigurđsson, 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., 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, 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, 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., 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., A.P.M.

COMPETING FINANCIAL INTERESTS

K. Stefansson, V. Steinthorsdottir, G.T., and U.T. are employed by deCODE Genetics/Amgen inc. I.B. and spouse own stock in GlaxoSmithKline and Incyte.

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 significant excess in 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 novel T2D susceptibility loci. Furthermore, we observed considerable improvements in 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 characterisation 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 ancestry[1–5], 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 groups[6]. However, European ancestry populations constitute only a subset of human genetic variation, and thus are insufficient to fully characterise T2D risk variants in other ethnic groups. Furthermore, the latest genome-wide genotyping arrays are less biased towards Europeans, and more recent T2D GWAS have been performed, with great success, in populations from other ancestry groups, including East Asians[7–12], South Asians[13,14], Mexicans and Mexican Americans[15], and African Americans[16]. These studies have provided initial evidence of overlap in T2D susceptibility loci between ancestry groups and 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 prior to human population migration out of Africa. Under such a model, we would expect to improve power to detect novel susceptibility loci for the disease, and enhance 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 due to less extensive LD than other ethnic groups[6], 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 novel 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

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) Consortium[5] (European ancestry; 12,171 cases and 56,862 controls); (ii) the Asian Genetic Epidemiology Network T2D (AGEN-T2D) Consortium[11] (East Asian ancestry; 6,952 cases and 11,865 controls); (iii) the South Asian T2D (SAT2D) Consortium[13] (South Asian ancestry; 5,561 cases and 14,458 controls); and (iv) the Mexican American T2D (MAT2D) Consortium[15] (Mexican and Mexican American ancestry; 1,804 cases and 779 controls). We obtained association summary statistics from the four available ethnic-specific meta-analyses, each imputed at up to 2.5 million autosomal SNPs from Phase II/III HapMap[22,23] to provide a uniform catalogue of common genetic variation, defined by minor allele frequency (MAF) of at least 5%, across ancestry groups (**Online Methods**). These association summary statistics were then combined across ancestry groups via trans-ethnic fixed-effects meta-analysis (**Online Methods**).

Directional consistency 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*-statistic from the trans-ethnic meta-analysis (**Online Methods**, 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 largest in European ancestry populations, 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 East Asian and European ancestry populations, 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 $p<2.2\times10^{-16}$). The strong evidence of homogeneity in allelic effects across ancestry groups at the majority of previously reported lead SNPs argues against the "synthetic association" hypothesis[24]. It is improbable that GWAS signals at most established T2D susceptibility loci reflect unobserved lower frequency causal alleles with larger effects because: (i) rare variants are unlikely to have arisen before human population migration out of Africa and thus are not expected to be widely shared across diverse populations[25]; 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 novel 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, defined as mapping within 500kb of the previously reported lead SNPs (**Online Methods**). First, we identified independent SNPs (separated by at least 500kb) with nominal evidence of association ($p\leq0.001$) 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 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 excess consistency in the direction of effect between ethnicities at independent SNPs demonstrating weaker evidence of T2D association ($0.001<p\leq0.01$) 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 East Asian, South Asian, and Mexican and Mexican American meta-analyses, in turn, 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 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 novel T2D susceptibility loci achieving genome-wide significance

The observations from our concordance analyses are consistent with a long tail of common T2D susceptibility variants, with effects which are decreasing in magnitude, but which are homogeneous across ancestry groups. Under such a model, we would expect these variants to be amenable to discovery via 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 500kb) with suggestive evidence of association ($p<10^{-5}$) at loci not previously reported for T2D susceptibility in any ancestry group (Supplementary Table 5, Supplementary Figure 1). By convention, we have labelled 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 T2D cases and 55,647 controls of European ancestry[5], genotyped with the Metabochip (**Online Methods**, 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 traits[26]. However, it provides relatively limited coverage of common genetic variation, genome-wide, with the result that the lead 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 (*SSR1/RREB1*, CEU $r^2=0.26$, $p=1.9\times10^{-6}$) and rs4275659 (*MPHOSPH9*, CEU $r^2=0.48$, $p=5.5\times10^{-6}$), which, nonetheless, demonstrated only marginally weaker association signals than the lead SNPs (*SSR1/RREB1*, rs9502570, $p=5.7\times10^{-7}$; *MPHOSPH9*, rs1727313, $p=1.2\times10^{-6}$). Given that these variants met our threshold for follow-up from the trans-ethnic meta-analysis, they were also considered for validation.

By combining association summary statistics from the trans-ethnic "discovery" and European ancestry "validation" meta-analyses, SNPs achieved genome-wide significance (combined meta-analysis $p<5\times10^{-8}$) at seven loci (Table 2, Figure 1). We observed no evidence of heterogeneity in allelic effects between discovery and validation stages of the combined meta-analysis (Supplementary Table 5). As expected, the novel loci are characterised by lead SNPs that are relatively common in all ethnicities, and have modest effects on T2D susceptibility which are homogeneous across ancestry groups (Supplementary Table 6). Adjustments for covariates were not harmonised 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 studies according to covariate adjustment (**Online Methods**, Supplementary Table 7). These data thus provide no evidence of bias in allelic effect estimates at lead SNPs at the novel loci, and suggest our results to be robust to variability in correction for potential confounders across studies.

The novel loci include SNPs mapping near *POU5F1/TCF19* in the major histocompatibility complex (MHC), a region of the genome that is essential to immune response. The MHC harbours HLA class II genes, which together account for approximately half 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, with the same risk allele for both diseases, at just two (*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], a late-age onset, more indolent form of the disease, which 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 LADA cases, the lack of association of T1D-risk variants suggests that rates of diagnostic misclassification of autoimmune diabetes were too modest to drive the T2D GWAS signal at the *POU5F1/TCF19* locus.

The novel loci also include SNPs mapping to *ARL15* and *SSR1/RREB1*, which have been previously implicated, at genome-wide significance, in regulation of fasting insulin (FI) and fasting glucose (FG), respectively[30]. The lead SNPs for T2D (rs702634) and FI (rs4865796) mapping to *ARL15* are closely correlated in European and East Asian ancestry populations (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 *SSR1/RREB1* locus (CEU and CHB+JPT $r^2<0.05$). The *ARL15* locus has also been associated with circulating adiponectin levels, an adipocyte-secreted protein that has anti-diabetic effects[31], 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 novel T2D susceptibility loci with metabolic phenotypes, we interrogated published European ancestry meta-analyses from the Meta-Analysis of Glycaemic and Insulin-related Consortium (MAGIC) Investigators[3,30], the

Genetic Investigation of ANthropometric Traits (GIANT) Consortium[32,33] and the Global Lipids Genetics Consortium[34], to evaluate the effect of T2D risk alleles on: glycaemic 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**, Supplementary Tables 9, 10 and 11). T2D risk alleles at *SSR1/RREB1* and *LPP* have features that indicate a primary role on susceptibility through beta-cell dysfunction: increased FG (p=1.0×10^{-5} and p=8.6×10^{-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=5.0×10^{-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 high-density lipoprotein cholesterol (p=0.022) and increased triglycerides (p=0.010), as expected, but also with *reduced* BMI (p=5.6×10^{-5}).

To identify the most promising functional candidate transcripts amongst those mapping to the novel 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}) 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=2.2×10^{-6}) at the *SSR1/RREB1* locus; *ABCB9* (liver, p=7.4×10^{-12}) and *SETD8* (lung, p<2.0×10^{-16}) at the *MPHOSPH9* locus; and *HCG27* (monocytes, p=1.3×10^{-69}) at the *POU5F1/TCF19* locus (Supplementary Table 12).

We also evaluated novel loci for potential functional mechanisms underlying T2D susceptibility (**Online Methods**). We identified variants in pilot data from the 1000 Genomes Project[25] that are in strong LD (CEU and CHB+JPT r^2>0.8) with the lead SNPs in the seven novel susceptibility loci for functional annotation. We identified a missense variant at the *POU5F1/TCF19* locus in *TCF19* (rs113581344, V211M; CEU r^2=0.96 and CHB+JPT r^2=0.80 with lead SNP rs3130501), although it is predicted to be tolerated by SIFT[35] (Supplementary Table 13). Lead SNPs in the novel susceptibility loci were also in strong LD with variants in the untranslated regions of *SSR1* (at the *SSR1/RREB1* locus) and *ABCB9*, *OGFOD2*, and *PITPNM2* (at the *MPHOSPH9* locus). Variants in strong LD with the lead SNPs at two of the novel susceptibility loci overlap regions of predicted regulatory function generated by the ENCODE Project[36] (Supplementary Figure 2). The lead SNP at the *LPP* locus maps to an enhancer region which 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) which overlaps a region of open chromatin activity in pancreatic islets and other cell types. This open chromatin site is in a region correlated with expression of *ELAVL4*, which has been demonstrated to regulate insulin translation in pancreatic beta cells[37], 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 variants[38], 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 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 MANTRA[39]. 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 localisation 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 sets"[41] of SNPs that are most likely to be causal based on their statistical evidence of association from the MANTRA meta-analysis. Credible sets can be interpreted in a similar way to confidence intervals in a frequentist statistical framework. For example, assuming that a locus harbours 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 is 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 magnitude of effects on T2D susceptibility. Consequently, we have not considered the European ancestry Metabochip validation studies in our fine-mapping analyses because SNP density on the array is too sparse, across the majority of T2D susceptibility loci, to allow high-quality imputation up to the Phase II/III HapMap reference panels utilised 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 characterised by independent common "index" SNPs, are relatively widespread at T2D susceptibility loci, for example *CDKN2A/B* and *KCNQ16*. 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 GCTA[42], 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: (i) the European ancestry only meta-analysis; and (ii) the trans-ethnic meta-analysis of European, East Asian, South Asian, and Mexican and

Mexican American ancestry groups. We focussed on ten 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 since they were typically characterised by large 99% credible sets in the European ancestry meta-analysis, and thus might provide an over-estimate of the improvement in fine-mapping resolution by combining GWAS across ancestry groups. Of the loci considered, only at *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.

The greatest enhancement in fine-mapping resolution after trans-ethnic meta-analysis was observed at the *JAZF1* locus, where the genomic interval covered by the 99% credible set was reduced from 76kb to just 16kb (Figure 2, Supplementary Figure 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 East Asian ancestry populations (CHB+JPT r^2<0.05 with rs864745). Amongst the variants retained in the 99% credible set after trans-ethnic meta-analysis, interrogation of predicted regulatory function from the ENCODE Project[36] revealed that rs1635852 maps to a region of open chromatin with enhancer activity, bound by several transcription factors. This SNP has been previously shown to have allelic differences in pancreatic islet enhancer activity[43], 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 (Figure 2, Supplementary Figure 3), from 35kb (four SNPs) on the basis of only European ancestry GWAS, to less than 1kb (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 European and East Asian ancestry groups (CEU and CHB+JPT r^2≥0.8 with rs13266634), suggesting that the improved fine-mapping resolution at this locus is more likely due to 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 others in the credible set, and is a non-synonymous variant, R325W, which plays an established functional role in T2D susceptibility[44].

Finally, we tested variants present in the 99% credible sets at the ten loci, on the basis of only the European ancestry GWAS 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, no such enrichment in either annotation category was observed 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 localisation of common causal SNPs within established T2D susceptibility loci, and 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 pass 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 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 resolution to be relatively robust to genotype quality at common variants.

DISCUSSION

We have identified seven novel 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, which contribute to the "missing heritability" of T2D susceptibility, in agreement with polygenic analyses in European ancestry GWAS[5,49]. The effects of these common variants are modest, but homogeneous across ancestry groups, and thus would 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 localisation due to reduced LD in these populations. Future releases of reference panels from the 1000 Genomes Project are anticipated to include 2,500 samples, including haplotypes of South Asian ancestry and wider representation of African descent populations. This panel will provide a comprehensive catalogue of genetic variation with MAF 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 characterisation of common variant loci contributing to complex traits, and emphasise an exciting opportunity to further our understanding of the biological mechanisms underlying human diseases across populations from diverse ethnicities.

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 individual 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%, (except MAF>5% in the Mexican and Mexican American ancestry GWAS due to 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 harmonised within or between consortia because of variation in individual study design and recorded non-genetic risk factors. The results of each GWAS were corrected for population structure with genomic control[50] (unless $\lambda_{GC}<1$). Association summary statistics from GWAS within each ancestry group were then combined via fixed-effects meta-analysis. The results of each ancestry 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” GWAS meta-analysis

Association summary statistics from each ancestry-specific meta-analysis were combined via 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 control[50] ($\lambda_{GC}=1.05$).

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-statistic[51] (Supplementary Table 3). Amongst 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 500kb up- and down-stream of the previously reported lead SNPs. We also excluded AT/GC SNPs to eliminate bias due to strand misalignment between ethnic-specific meta-analyses. Amongst the remaining SNPs, we selected an independent subset with nominal evidence of association ($p\leq0.001$) with T2D from the European ancestry meta-analysis, separated by at least 500kb. 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: $0.001<p\leq0.01$; $0.01<p\leq0.5$; and $0.5<p\leq1$ (Table 1). Finally, we 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" meta-analysis

The previously published validation meta-analysis consisted of 21,491 cases and 55,647 controls of European ancestry from the DIAGRAM Consortium[5], all genotyped with the Metabochip[26] (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-analysis[13]. Full details of the samples and methods employed in the validation meta-analysis are presented in the DIAGRAM Consortium paper[5]. 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 via 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 novel 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, 16 SNPs were genotyped directly on Metabochip, and 10 more had a proxy (CEU and CHB+JPT HapMap $r^2\geq0.2$). For these 26 SNPs, association summary statistics from the discovery and validation meta-analyses were combined via 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*-statistic[51].

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 via 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 via fixed-effects inverse-variance weighted meta-analysis, but did not require subsequent correction for genomic control (λ_{QT}=1.00). We then tested for heterogeneity in allelic effects between these two sets of studies by means of Cochran's *Q*-statistic[51] (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 loci[27] (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-DR4* (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 novel T2D susceptibility loci in published European ancestry GWAS meta-analyses of metabolic phenotypes: glycaemic traits[3,30], anthropometric measures[32,33], and plasma lipid concentrations[34]. 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. Finally, we 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 novel susceptibility loci in multiple tissues. Details of these resources are summarised 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 Project[25] (CEU and CHB+JPT) to assess coincidence of the signals.

Functional annotation

We identified variants in pilot data from the 1000 Genomes Project[25] that are in strong LD (CEU and CHB+JPT r^2>0.8) with the lead SNPs in the novel susceptibility loci for

functional annotation. Identified non-synonymous variants were interrogated for likely downstream functional consequences using SIFT[35] (Supplementary Table 13). Variants were also assessed for overlap with regions of predicted regulatory function generated by the ENCODE Project[36] 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 Figure 2).

Fine-mapping analyses

We used MANTRA[39] to fine-map T2D susceptibility loci on the basis of association summary statistics from: (i) the meta-analysis of European ancestry GWAS only[5]; and (ii) 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, defined by pair-wise genome-wide mean allele frequency differences (Supplementary Figure 4). Evidence in favour 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 500kb up- and down-stream of the lead SNP from the trans-ethnic meta-analysis). We then calculated the posterior probability that the *j*th SNP is causal, amongst those reported in the meta-analysis, by:

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

In this expression, BF_j denotes the BF in favour of association of the *j*th SNP, and the summation in the denominator is over all variants passing QC across the locus[41]. A 99% credible set of variants was then constructed by: (i) ranking all SNPs according to their BF; and (ii) 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 Project[36]. We performed 1,000 permutations by shifting the location of the annotation sites a random distance within 100kb, and recalculated the overlap to obtain empirical *p*-values for enrichment in each annotation category.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding for the research undertaken in this study has been received from: 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.

References

1. 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. 2008; 40:638–45. [PubMed: 18372903]
2. Kong A, et al. Parental origin of sequence variants associated with complex diseases. Nature. 2009; 462:868–74. [PubMed: 20016592]
3. Dupuis J, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010; 42:105–16. [PubMed: 20081858]
4. Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large scale association analysis. Nat Genet. 2010; 42:579–89. [PubMed: 20581827]
5. Morris AP, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012; 44:981–90. [PubMed: 22885922]
6. Rosenberg NA, et al. Genome-wide association studies in diverse populations. Nat Rev Genet. 2010; 11:356–66. [PubMed: 20395969]
7. Qi L, et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Hum Mol Genet. 2010; 19:2706–15. [PubMed: 20418489]
8. Tsai F-J, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet. 2010; 6:e1000847. [PubMed: 20174558]
9. Shu XO, et al. Identification of new genetic risk variants for type 2 diabetes. PLoS Genet. 2010; 6:e1001127. [PubMed: 20862305]
10. 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. 2010; 42:864–8. [PubMed: 20818381]
11. Cho YS, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. Nat Genet. 2012; 44:67–72. [PubMed: 22158537]
12. Li H, et al. A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. Diabetes. 2013; 62:91–8.
13. Kooner JS, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet. 2011; 43:984–9. [PubMed: 21874001]
14. Tabassum R, et al. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. Diabetes. 2013; 62:977–86. [PubMed: 23209189]
15. Parra EJ, 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. 2011; 54:2038–46. [PubMed: 21573907]
16. Palmer ND, et al. A genome-wide association search for type 2 diabetes genes in African Americans. PLoS One. 2012; 7:e29202. [PubMed: 22238593]
17. Waters KM, et al. Consistent association of type 2 diabetes risk variants found in Europeans in diverse racial and ethnic groups. PLoS Genet. 2010; 6:e1001078. [PubMed: 20865176]
18. Saxena R, et al. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. Am J Hum Genet. 2012; 90:410–25. [PubMed: 22325160]
19. Cooper RS, Tayo B, Zhu X. Genome-wide association studies: implications for multi-ethnic samples. Hum Mol Genet. 2008; 17:R151–R155. [PubMed: 18852204]

20. Zaitlen N, Pasaniuc B, Gur T, Ziv E, Halperin E. Leveraging genetic variability across populations for the identification of causal variants. Am J Hum Genet. 2010; 86:23–33. [PubMed: 20085711]
21. Fanceschini N, et al. Discovery and fine-mapping of serum protein loci through transethnic meta-analysis. Am J Hum Genet. 2012; 91:744–53. [PubMed: 23022100]
22. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–61. [PubMed: 17943122]
23. The International HapMap Consortium. Integrating common and rare genetic variation in diverse human populations. Nature. 2010; 467:52–8. [PubMed: 20811451]
24. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]
25. The 1000 Genomes Project Consortium. A map of human genome variation from population scale sequencing. Nature. 2010; 467:1061–73. [PubMed: 20981092]
26. Voight BF, et al. The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. 2012; 8:e1002793. [PubMed: 22876189]
27. Bradfield JP, et al. A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet. 2011; 7:e1002293. [PubMed: 21980299]
28. Cervin C, et al. Genetic similarities between latent autoimmune diabetes in adults, type 1 diabetes, and type 2 diabetes. Diabetes. 2008; 57:1433–7. [PubMed: 18310307]
29. Grant SF, 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. 2010; 31:183–93. [PubMed: 20007922]
30. Scott RA, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet. 2012; 44:991–1005. [PubMed: 22885924]
31. Richards JN, et al. A genome-wide association study reveals variants in ARL15 that influence adiponectin levels. PLoS Genet. 2009; 5:e1000768. [PubMed: 20011104]
32. Speliotes EK, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010; 42:937–48. [PubMed: 20935630]
33. Heid IM, 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. 2010; 42:949–60. [PubMed: 20935629]
34. Teslovich TM, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010; 466:707–13. [PubMed: 20686565]
35. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003; 31:3812–4. [PubMed: 12824425]
36. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
37. Lee EK, et al. RNA-binding protein HuD controls insulin translation. Mol Cell. 2012; 45:826–35. [PubMed: 22387028]
38. Trynka G, et al. Chromatin marks identify critical cell types for fine-mapping complex trait variants. Nat Genet. 2013; 45:124–30. [PubMed: 23263488]
39. Morris AP. Transethnic meta-analysis of genomewide association studies. Genet Epidemiol. 2011; 35:809–22. [PubMed: 22125221]
40. Wang X, et al. Comparing methods for performing trans-ethnic meta-analysis of genome-wide association studies. Hum Mol Genet. 2013; 22:2302–11.
41. Maller JB, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat Genet. 2012; 44:1294–301. [PubMed: 23104008]
42. Yang J, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012; 44:369–75. [PubMed: 22426310]
43. Fogarty MP, Panhuis TM, Vadlamudi S, Buchkovich ML, Mohlke KL. Allele-specific transcriptional activity at type 2 diabetes-associated single nucleotide polymorphisms in regions of pancreatic islet open chromatin at the JAZF1 locus. Diabetes. 2013; 62:1756–62. [PubMed: 23328127]

44. Nicolson TJ, 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. 2009; 58:2070–83. [PubMed: 19542200]

45. The 1000 Genomes Project Consortium . An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]

46. Sung YJ, Wang L, Rankinen T, Bouchard C, Rao DC. Performance of genotype imputations using data from the 1000 Genomes Project. Hum Hered. 2012; 73:18–25. [PubMed: 22212296]

47. Zheng HF, Ladouceur M, Greenwood CM, Richards JB. Effect of genome-wide genotyping and reference panels on rare variant imputation. J Genet Genomics. 2012; 39:545–50. [PubMed: 23089364]

48. Nelson SC, et al. Imputation-based genomic coverage assessments of current human genotyping arrays. G3 (Bethesda). 2013; 3:1795–807. [PubMed: 23979933]

49. Stahl EA, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet. 2012; 44:483–9. [PubMed: 22446960]

50. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]

51. Ioannidis J, et al. Heterogeneity in meta-analyses of genome-wide association investigations. PLoS One. 2007; 2:e0000841.

Appendix

Anubha Mahajan[1,220], Min Jin Go[2,220], Weihua Zhang[3,220], Jennifer E Below[4,220], Kyle J Gaulton[1,220], Teresa Ferreira[1], Momoko Horikoshi[1,5], Andrew D Johnson[6], Maggie CY Ng[7,8], Inga Prokopenko[1,5,9], Danish Saleheen[10,11], Xu Wang[12], Eleftheria Zeggini[13], Goncalo R Abecasis[14], Linda S Adair[15], Peter Almgren[16], Mustafa Atalay[17], Tin Aung[18,19], Damiano Baldassarre[20,21], Beverley Balkau[22,23], Yuqian Bao[24], Anthony H Barnett[25,26], Ines Barroso[13,27,28], Abdul Basit[29], Latonya F Been[30], John Beilby[31,32,33],

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK.
[220]These authors contributed equally to this work.
[2]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Cheongwon-gun, Gangoe-myeon, Yeonje-ri, Korea.
[3]Epidemiology and Biostatistics, Imperial College London, London, UK.
[4]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA.
[5]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, OX3 7LJ, UK.
[6]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA 01702, USA.
[7]Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
[8]Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
[9]Genomics of Common Disease, Imperial College London, Hammersmith Hospital, W12 0NN, London, UK.
[10]Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK.
[11]Center for Non-Communicable Diseases Pakistan, Karachi, Pakistan.
[12]Department of Epidemiology and Public Health, National University of Singapore, Singapore, Singapore.
[13]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.
[14]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109-2029, USA.
[15]Department of Nutrition, University of North Carolina, Chapel Hill, North Carolina, USA.
[16]Lund University Diabetes Centre, Department of Clinical Science Malmö, Scania University Hospital, Lund University, S-20502 Malmö, Sweden.
[17]Institute of Biomedicine, Physiology, University of Eastern Finland, Kuopio Campus, Kuopio, Finland.
[18]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore.
[19]Department of Ophthalmology, National University of Singapore, Singapore, Singapore.
[20]Centro Cardiologico Monzino, IRCCS, via Carlo Parea 4, 20138, Milan, Italy.
[21]Department of Pharmacological Sciences, University of Milan, via Balzaretti 9, 20133 Milan, Italy.
[22]INSERM CESP U1018, F-94807 Villejuif, France.
[23]University Paris Sud 11, UMRS 1018, F-94807, Villejuif, France.
[24]Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Department of Endocrinology and Metabolism, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China.
[25]College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK.
[26]BioMedical Research Centre, Heart of England NHS Foundation Trust, Birmingham, UK.

Graeme I Bell[34,35], Rafn Benediktsson[36,37], Richard N Bergman[38], Bernhard O Boehm[39,40], Eric Boerwinkle[41,42], Lori L Bonnycastle[43], Noël Burtt[44], Qiuyin Cai[45], Harry Campbell[46,47], Jason Carey[44], Stephane Cauchi[48], Mark Caulfield[49], Juliana CN Chan[50], Li-Ching Chang[51], Tien-Jyun Chang[52], Yi-Cheng Chang[52], Guillaume Charpentier[53], Chien-Hsiun Chen[51,54], Han Chen[55], Yuan-Tsong Chen[51], Kee-Seng Chia[12,56], Manickam Chidambaram[57], Peter S Chines[43], Nam H Cho[58], Young Min Cho[59], Lee-Ming Chuang[52,60], Francis S Collins[43], Marilyn C Cornelis[61], David J Couper[62], Andrew T Crenshaw[44], Rob M van Dam[61,63], John Danesh[10], Debashish Das[64], Ulf de Faire[65], George Dedoussis[66], Panos Deloukas[13], Antigone S Dimas[1,67,68], Christian Dina[69,70,71], Alex SF Doney[72,73], Peter J Donnelly[1,74], Mozhgan Dorkhan[16], Cornelia

[27]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Box 289, Addenbrooke's Hospital, Cambridge, CB2 OQQ, UK.
[28]NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Box 289, Addenbrooke's Hospital, Cambridge, CB2 OQQ, UK.
[29]Baqai Institute of Diabetology and Endocrinology, Karachi, Pakistan.
[30]Department of Pediatrics, Section of Genetics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA.
[31]Busselton Population Medical Research Institute, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
[32]PathWest Laboratory Medicine of Western Australia, QEII Medical Centre, Nedlands, WA 6009, Australia.
[33]School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, WA 6009, Australia.
[34]Department of Medicine, University of Chicago, Chicago, IL, 60637, USA.
[35]Department of Human Genetics, University of Chicago, Chicago, IL, 60637, USA.
[36]Faculty of Medicine, University of Iceland, 101 Reykjavík, Iceland.
[37]Landspitali University Hospital, 101 Reykjavík, Iceland.
[38]Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA.
[39]Division of Endocrinology and Diabetes, Department of Internal Medicine, University Medical Centre Ulm, Ulm, Germany.
[40]LKC School of Medicine, Nanyang Technological University, Singapore, Singapore, and Imperial College London, London, UK.
[41]Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
[42]Human Genome Sequencing Center at Baylor College of Medicine, Houston, TX 77030, USA.
[43]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA.
[44]Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, MA 02142, USA.
[45]Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA.
[46]Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, UK.
[47]MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, UK.
[48]CNRS-UMR-8199, Institute of Biology and Lille 2 University, Pasteur Institute, F-59019 Lille, France.
[49]Clinical 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.
[50]Department of Medicine and Therapeutics, Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China.
[51]Institute of Biomedical Sciences, Academia Sinica, Nankang, Taipei, Taiwan.
[52]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan.
[53]Endocrinology-Diabetology Unit, Corbeil-Essonnes Hospital, F-91100 Corbeil-Essonnes, France.
[54]School of Chinese Medicine, China Medical University, Taichung, Taiwan.
[55]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA.
[56]Centre for Molecular Epidemiology, National University of Singapore, Singapore, Singapore.
[57]Department of Molecular Genetics, Madras Diabetes Research Foundation–Indian Council of Medical Research (ICMR) Advanced Centre for Genomics of Diabetes, Chennai, India.
[58]Department of Preventive Medicine, Ajou University School of Medicine, Suwon, Korea.
[59]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea.
[60]Graduate Institute of Clinical Medicine, National Taiwan University School of Medicine, Taipei, Taiwan.
[61]Department of Nutrition and Epidemiology, Harvard School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA.
[62]Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[63]Saw Swee Hock School of Public Health, National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077.
[64]Ealing Hospital National Health Service (NHS) Trust, Middlesex, UK.
[65]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
[66]Department of Dietetics-Nutrition, Harokopio University, 70 El. Venizelou Str, Athens, Greece.
[67]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva 1211, Switzerland.
[68]Biomedical Sciences Research Center Al. Fleming, 16672 Vari, Greece.
[69]Inserm UMR 1087, 44007 Nantes, France.
[70]CNRS UMR 6291, 44007 Nantes, France.

van Duijn[75,76], Josée Dupuis[6,55], Sarah Edkins[13], Paul Elliott[3,77], Valur Emilsson[78], Raimund Erbel[79], Johan G Eriksson[80,81,82,83], Jorge Escobedo[84], Tonu Esko[85,86,87,88], Elodie Eury[48], Jose C Florez[87,89,90,91], Pierre Fontanillas[44], Nita G Forouhi[92], Tom Forsen[81,93], Caroline Fox[6,94], Ross M Fraser[46], Timothy M Frayling[95], Philippe Froguel[9,48], Philippe Frossard[11], Yutang Gao[96], Karl Gertow[97,98], Christian Gieger[99], Bruna Gigante[65], Harald Grallert[100,101,102,103], George B Grant[44], Leif C Groop[16], Christopher J Groves[5], Elin Grundberg[104], Candace Guiducci[44], Anders Hamsten[97,98], Bok-Ghee Han[2], Kazuo Hara[105], Neelam Hassanali[5], Andrew T Hattersley[106], Caroline Hayward[47], Asa K Hedman[1], Christian Herder[107], Albert Hofman[75], Oddgeir L Holmen[108], Kees Hovingh[109], Astradur B Hreidarsson[37], Cheng Hu[24], Frank B Hu[61,110],

[71]Nantes University, 44007 Nantes, France.
[72]Diabetes Research Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee DD1 9SY, UK.
[73]Pharmacogenomics Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee DD1 9SY, UK.
[74]Department of Statistics, University of Oxford, 1 South Parks Road Oxford, OX1 3TG, UK.
[75]Department of Epidemiology, Erasmus University Medical Center, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands.
[76]Netherland Genomics Initiative, Netherlands Consortium for Healthy Ageing and Centre for Medical Systems Biology, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands.
[77]Medical Research Council (MRC)-Health Protection Agency (HPA) Centre for Environment and Health, Imperial College London, London, UK.
[78]Icelandic Heart Association, 201 Kopavogur, Iceland.
[79]Clinic of Cardiology, West German Heart Centre, University Hospital of Essen, University Duisburg-Essen, 45122 Essen, Germany.
[80]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki FIN-00271, Finland.
[81]Department of General Practice and Primary Health Care, University of Helsinki, 00014 Helsinki, Finland.
[82]Unit of General Practice, Helsinki University General Hospital, Helsinki, Finland.
[83]Folkhälsan Research Center, FIN-00014 Helsinki, Finland.
[84]Unidad de Investigacion en Epidemiologia Clinica, Hospital General Regional I, Dr. Carlos MacGregor, IMSS, Mexico City, Mexico.
[85]Estonian Genome Center, University of Tartu, Tartu, Estonia.
[86]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.
[87]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA.
[88]Division of Genetics and Endocrinology, Children's Hospital, Boston, MA 02115, USA.
[89]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA.
[90]Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA.
[91]Diabetes Research Center, Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA.
[92]MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK.
[93]Vaasa Health Care Centre, 65100 Vaasa, Finland.
[94]Division of Endocrinology and Metabolism, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.
[95]Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Magdalen Road, Exeter EX1 2LU, UK.
[96]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China.
[97]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
[98]Center for Molecular Medicine, Karolinska University Hospital Solna, Stockholm, Sweden.
[99]Institute of Genetic Epidemiology, Helmholtz Zentrum Muenchen, 85764 Neuherberg, Germany.
[100]Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, 85764 Neuherberg, Germany.
[101]Clinical Cooperation Group Diabetes, Ludwig-Maximilians-Universität Muenchen and Helmholtz Zentrum Muenchen, Germany.
[102]Clinical Cooperation Group Nutrigenomics and Type 2 Diabetes, Technical University Muenchen and Helmholtz Zentrum Muenchen, Germany.
[103]German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany.
[104]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
[105]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.
[106]Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Barrack Road, Exeter EX2 5DW, UK.
[107]Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.
[108]HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway.
[109]Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, PO-BOX 22660, 1100DD, Amsterdam, The Netherlands.
[110]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Ave, Boston, MA 02115, USA.

Jennie Hui[31,32,33,111], Steve E Humphries[112], Sarah E Hunt[13], David J Hunter[61,110,113], Kristian Hveem[108], Zafar I Hydrie[29], Hiroshi Ikegami[114], Thomas Illig[100,115], Erik Ingelsson[1,116], Muhammed Islam[117], Bo Isomaa[83,118], Anne U Jackson[14], Tazeen Jafar[117,119], Alan James[31,120,121], Weiping Jia[24], Karl-Heinz Jöckel[122], Anna Jonsson[16], Jeremy BM Jowett[123], Takashi Kadowaki[105], Hyun Min Kang[14], Stavroula Kanoni[13], Wen Hong L Kao[124], Sekar Kathiresan[44,90,125], Norihiro Kato[126], Prasad Katulanda[5,127], Sirkka M Keinanen-Kiukaanniemi[128,129], Ann M Kelly[25,26], Hassan Khan[10], Kay-Tee Khaw[10], Chiea-Chuen Khor[18,56,130], Hyung-Lae Kim[131], Sangsoo Kim[132], Young Jin Kim[2], Leena Kinnunen[133], Norman Klopp[100,115], Augustine Kong[134], Eeva Korpi-Hyövälti[135], Sudhir Kowlessur[136], Peter Kraft[61,113], Jasmina Kravic[16], Malene M Kristensen[123], S Krithika[137], Ashish Kumar[1], Jesus Kumate[138], Johanna Kuusisto[139], Soo Heon Kwak[59], Markku Laakso[139], Vasiliki Lagou[1], Timo A Lakka[17,140], Claudia Langenberg[92], Cordelia Langford[13], Robert Lawrence[141], Karin Leander[65], Jen-Mai Lee[12], Nanette R Lee[142], Man Li[124], Xinzhong Li[143], Yun Li[144,145], Junbin Liang[146], Samuel Liju[57], Wei-Yen Lim[56], Lars Lind[147], Cecilia M Lindgren[1,44], Eero Lindholm[16], Ching-Ti Liu[55], Jian Jun Liu[130], Stéphane Lobbens[48], Jirong Long[45], Ruth JF Loos[92,148,149,150],

[111]School of Population Health, The University of Western Australia, Nedlands, WA 6009, Australia.
[112]Institute of Cardiovascular Science, University College London, 5 University Street, London WC1E 6JJ, UK.
[113]Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA.
[114]Department of Endocrinology, Metabolism and Diabetes, Kinki University School of Medicine, Osaka-sayama, Osaka, Japan.
[115]Hannover Unified Biobank, Hannover Medical School, 30625 Hannover, Germany.
[116]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University Hospital, SE-751 85 Uppsala, Sweden.
[117]Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan.
[118]Department of Social Services and Health Care, 68601 Jakobstad, Finland.
[119]Department of Medicine, Aga Khan University, Karachi, Pakistan.
[120]Department of Pulmonary Physiology and Sleep Medicine, West Australian Sleep Disorders Research Institute, Queen Elizabeth II Medical Centre, Hospital Avenue, Nedlands WA 6009, Australia.
[121]School of Medicine and Pharmacology, University of Western Australia, Nedlands WA 6009, Australia.
[122]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, Germany.
[123]Baker IDI Heart and Diabetes Institute, Melbourne, Victoria, Australia.
[124]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21287, USA.
[125]Cardiovascular Research Center, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA.
[126]Research Institute, National Center for Global Health and Medicine, Shinjuku-ku, Tokyo, Japan.
[127]Diabetes Research Unit, Department of Clinical Medicine, University of Colombo, Colombo, Sri Lanka.
[128]Faculty of Medicine, Institute of Health Sciences, University of Oulu, Oulu, Finland.
[129]Unit of General Practice, Oulu University Hospital, Oulu, Finland.
[130]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore.
[131]Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Korea.
[132]School of Systems Biomedical Science, Soongsil University, Dongjak-gu, Seoul, Korea.
[133]Diabetes Prevention Unit, National Institute for Health and Welfare, 00271 Helsinki, Finland.
[134]deCODE Genetics, 101 Reykjavik, Iceland.
[135]South Ostrobothnia Central Hospital, 60220 Seinäjoki, Finland.
[136]Ministry of Health, Port Louis, Mauritius.
[137]Department of Anthropology, University of Toronto at Mississauga, 3359 Mississauga Road North, Mississauga, ON L5L 1C6, Canada.
[138]Fundacion IMSS, Av. Paseo de la Reforma 476, Mz. Poniente, Col. Juarez, Deleg. Cuauhtemoc, C.P. 06600, Mexico City, Mexico.
[139]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, 70210 Kuopio, Finland.
[140]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland.
[141]Centre for Genetic Epidemiology and Biostatistics, The University of Western Australia, Nedlands, WA 6009, Australia.
[142]Office of Population Studies Foundation Inc., University of San Carlos, Cebu City, Philippines.
[143]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK.
[144]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA.
[145]Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, USA.
[146]Beijing Genomics Institute, Shenzhen, China.
[147]Department of Medical Sciences, Uppsala University, Akademiska Sjukhuset, SE-751 85 Uppsala, Sweden.
[148]Charles R. Bronfman Institute for Personalized Medicine, Mount Sinai School of Medicine, New York, NY, USA.

Wei Lu[151], Jian'an Luan[92], Valeriya Lyssenko[16], Ronald CW Ma[50], Shiro Maeda[60], Reedik Mägi[85], Satu Männistö[80], David R Matthews[5], James B Meigs[89,152], Olle Melander[16], Andres Metspalu[85,86], Julia Meyer[99], Ghazala Mirza[1], Evelin Mihailov[85], Susanne Moebus[122], Viswanathan Mohan[57,153], Karen L Mohlke[144], Andrew D Morris[72,73], Thomas W Mühleisen[154,155], Martina Müller-Nurasyid[99,156,157], Bill Musk[31,111,121,158], Jiro Nakamura[159], Eitaro Nakashima[159,160], Pau Navarro[47], Peng-Keat Ng[12], Alexandra C Nica[67], Peter M Nilsson[16], Inger Njølstad[161], Markus M Nöthen[154,155], Keizo Ohnaka[162], Twee Hee Ong[12], Katharine R Owen[5,163], Colin NA Palmer[72,73], James S Pankow[164], Kyong Soo Park[59,165], Melissa Parkin[44], Sonali Pechlivanis[122], Nancy L Pedersen[166], Leena Peltonen[13,44,80,167,219], John RB Perry[1,95], Annette Peters[168], Janani M Pinidiyapathirage[169], Carl GP Platou[108,170], Simon Potter[13], Jackie F Price[46], Lu Qi[61,110], Venkatesan Radha[57], Loukianos Rallidis[171], Asif Rasheed[11], Wolfgang Rathmann[172], Rainer Rauramaa[140,173], Soumya Raychaudhuri[44,174,175], N William Rayner[1,5], Simon D Rees[25,26], Emil Rehnberg[166], Samuli Ripatti[13,80,167], Neil Robertson[1,5], Michael Roden[107,176,177], Elizabeth J Rossin[44,87,178,179], Igor Rudan[46], Denis Rybin[180], Timo E Saaristo[181,182], Veikko Salomaa[80], Juha Saltevo[183], Maria

[149]Child Health and Development Institute, Mount Sinai School of Medicine, New York, NY 10029, USA.
[150]Department of Preventive Medicine, Mount Sinai School of Medicine, New York, NY 10029, USA.
[151]Shanghai Institute of Preventive Medicine, Shanghai, China.
[152]General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA.
[153]Dr Mohan's Diabetes Specialties Centre, Chennai, India.
[154]Institute of Human Genetics, University of Bonn, 53127 Bonn, Germany.
[155]Department of Genomics, Life & Brain Center, University of Bonn, 53127 Bonn, Germany.
[156]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany.
[157]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-University, Munich, Germany.
[158]Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
[159]Division of Endocrinology and Diabetes, Department of Internal Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan.
[160]Department of Diabetes and Endocrinology, Chubu Rosai Hospital, Nagoya, Japan.
[161]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, N-9037 Tromsø, Norway.
[162]Department of Geriatric Medicine, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan.
[163]Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Old Road Headington, Oxford, OX3 7LJ, UK.
[164]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 55454, USA.
[165]World Class University Program, Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology and College of Medicine, Seoul National University, Seoul, Korea.
[166]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, SE-171 77 Stockholm, Sweden.
[167]Institute for Molecular Medicine Finland (FIMM), Helsinki FIN-00014, Finland.
[219]Deceased.
[168]Institute of Epidemiology II, Helmholtz Zentrum Muenchen, 85764 Neuherberg, Germany.
[169]Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka
[170]Department of Internal Medicine, Levanger Hospital, Nord-Trøndelag Health Trust, B-7600 Levanger, Norway.
[171]University General Hospital Attikon, 1 Rimini, 12462 Chaidari, Athens, Greece.
[172]Institute of Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.
[173]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland.
[174]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.
[175]Partners Center for Personalized Genomic Medicine, Boston, MA 02115, USA.
[176]Department of Endocrinology and Diabetology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.
[177]Department of Metabolic Diseases, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.
[178]Health Science and Technology MD Program, Harvard University and Massachusetts Institute of Technology, Boston, Massachusetts, USA.
[179]Harvard Biological and Biomedical Sciences Program, Harvard University, Boston, Massachusetts, USA.
[180]Boston University Data Coordinating Center, Boston, Massachusetts, USA.
[181]Finnish Diabetes Association, Kirjoniementie 15, 33680, Tampere, Finland.
[182]Pirkanmaa Hospital District, Tampere, Finland.
[183]Department of Medicine, Central Finland Central Hospital, 40620 Jyväskylä, Finland.

Samuel[11], Dharambir K Sanghera[30], Jouko Saramies[184], James Scott[143], Laura J Scott[14], Robert A Scott[92], Ayellet V Segrè[44,89,90], Joban Sehmi[64,143], Bengt Sennblad[97,98], Nabi Shah[11], Sonia Shah[185], A Samad Shera[186], Xiao Ou Shu[45], Alan R Shuldiner[187,188,189], Gunnar Sigurđsson[37,78], Eric Sijbrands[190], Angela Silveira[97,98], Xueling Sim[14,56], Suthesh Sivapalaratnam[109], Kerrin S Small[13,104], Wing Yee So[50], Alena Stančáková[139], Kari Stefansson[36,134], Gerald Steinbach[191], Valgerdur Steinthorsdottir[134], Kathleen Stirrups[13], Rona J Strawbridge[97,98], Heather M Stringham[14], Qi Sun[61,110], Chen Suo[56], Ann-Christine Syvänen[192], Ryoichi Takayanagi[193], Fumihiko Takeuchi[126], Wan Ting Tay[18], Tanya M Teslovich[14], Barbara Thorand[168], Gudmar Thorleifsson[134], Unnur Thorsteinsdottir[36,134], Emmi Tikkanen[80,167], Joseph Trakalo[1], Elena Tremoli[20,21], Mieke D Trip[109], Fuu Jen Tsai[54], Tiinamaija Tuomi[83,194], Jaakko Tuomilehto[133,195,196,197], Andre G Uitterlinden[75,76,190], Adan Valladares-Salgado[198], Sailaja Vedantam[87,88], Fabrizio Veglia[20], Benjamin F Voight[44,199], Congrong Wang[24], Nicholas J Wareham[92], Roman Wennauer[190], Ananda R Wickremasinghe[169], Tom Wilsgaard[161], James F Wilson[46,47], Steven Wiltshire[1,219], Wendy Winckler[44], Tien Yin Wong[18,19,200], Andrew R Wood[95], Jer-Yuarn Wu[51,54], Ying Wu[144], Ken Yamamoto[201], Toshimasa Yamauchi[105], Mingyu Yang[146], Loic Yengo[48,202], Mitsuhiro Yokota[203], Robin Young[10], Delilah Zabaneh[185], Fan Zhang[146], Rong Zhang[24], Wei Zheng[45], Paul Z Zimmet[123], David Altshuler[44,89,90,204,205,206,221], 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

[184]South Karelia Central Hospital, Lappeenranta, Finland.
[185]UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK.
[186]Diabetic Association Pakistan, Karachi, Pakistan.
[187]Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, MD 21201, USA.
[188]Geriatric Research Education and Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD 21201, USA.
[189]Program in Personalised and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA.
[190]Department of Internal Medicine, Erasmus University Medical Center, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands.
[191]Department of Clinical Chemistry and Central Laboratory, University of Ulm, Ulm, Germany.
[192]Molecular Medicine, Department of Medical Sciences, Uppsala University, SE-751 85 Uppsala, Sweden.
[193]Department of Internal Medicine and Bioregulatory Science, Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan.
[194]Department of Medicine, Helsinki University Hospital, University of Helsinki, 000290 HUS Helsinki, Finland.
[195]Instituto de Investigacion Sanitaria del Hospital Universario LaPaz (IdiPAZ), Madrid, Spain.
[196]Centre for Vascular Prevention, Danube-University Krems, 3500 Krems, Austria.
[197]Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia.
[198]Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, IMSS, Av. Cuauhtemoc 330, Col. Doctores, C.P. 06720, Mexico City, Mexico.
[199]University of Pennsylvania - Perelman School of Medicine, Department of Pharmacology, Philadelphia PA 19104, USA.
[200]Centre for Eye Research Australia, University of Melbourne, East Melbourne, Victoria, Australia.
[201]Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka, Japan.
[202]University Lille 1, Laboratory of Mathematics, CNRS-UMR 8524, MODAL team, INRIA Lille Nord-Europe.
[203]Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya, Japan.
[204]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
[205]Department of Molecular Biology, Harvard Medical School, Boston, MA 02115, USA.
[206]Diabetes Unit, Massachusetts General Hospital, Boston, MA 02144, USA.
[221]These authors jointly directed this work.
[207]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina;
[208]Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
[209]Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, Korea.
[210]Human Genetics Center, University of Texas Health Science Center at Houston, P.O. Box 20186, Houston, TX 77225, USA.
[211]Imperial College Healthcare NHS Trust, London, UK.
[212]Institute for Human Genetics, University of California, San Francisco, California, USA.

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], and Andrew P Morris[1,218,221]

[213]Blood Systems Research Institute, San Francisco, California, USA.
[214]Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore.
[215]Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore.
[216]Department of Medicine, National University of Singapore, Singapore, Singapore.
[217]Duke-National University of Singapore Graduate Medical School, Singapore, Singapore.
[218]Department of Biostatistics, University of Liverpool, Liverpool, L69 3GA, UK.

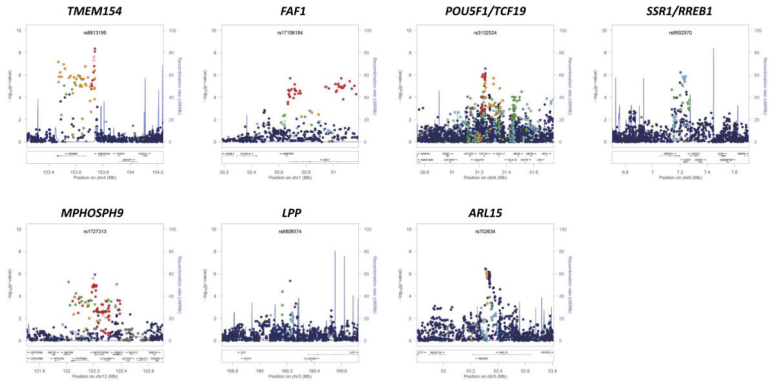

Figure 1. Signal plots of the trans-ethnic "discovery" GWAS meta-analysis for novel T2D susceptibility loci

The trans-ethnic meta-analysis comprises 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 their *p*-value (on a $-\log_{10}$ scale) as a function of genomic position (NCBI Build 36). In each panel, the lead SNP is represented by the purple symbol. The colour coding of all other SNPs indicates LD with the lead SNP (estimated by CEU r^2 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 $r^2 < 0.2$; grey r^2 unknown. The shape of the plotting symbol corresponds to the annotation of the SNP: upward triangle for framestop or splice; downward triangle for non-synonymous; square for synonymous or UTR; and circle for intronic or non-coding. Recombination rates are estimated from Phase II HapMap and gene annotations are taken from the University of California Santa Cruz genome browser.

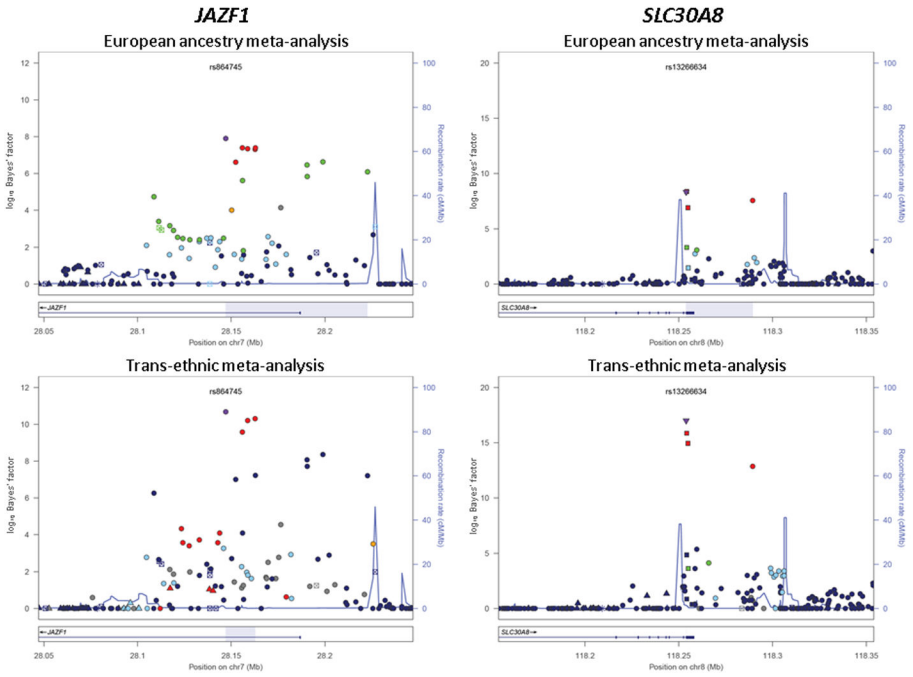

Figure 2. Signal plots presenting 99% credible sets of SNPs at the *JAZF1* and *SLC30A8* loci

The credible sets were constructed on the basis of: (i) the meta-analysis of European ancestry GWAS only (12,171 cases and 56,862 controls); and (ii) 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 panel, each point represents a SNP passing quality control in the MANTRA analysis, plotted with their 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 colour coding of all other SNPs indicates LD with the lead SNP (estimated by Phase II HapMap CEU r^2 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 $r^2 < 0.2$; grey r^2 unknown. The shape of the plotting symbol corresponds to the annotation of the SNP: upward triangle for framestop or splice; downward triangle for non-synonymous; square for synonymous or UTR; and circle for intronic or non-coding. Recombination rates are estimated from Phase II HapMap and gene annotations are taken from the University of California Santa Cruz genome browser. The genomic region covered by the 99% credible set is highlighted in grey.

Table 1

Concordance in the direction of effect of T2D risk alleles identified in a meta-analysis of GWAS of European ancestry (12,171 cases and 56,862 controls) with those from 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, defined as mapping within 500kb of the previously reported lead SNP.

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*-value	Concordant SNPs/Total SNPs	%	Binomial test *p*-value	Concordant SNPs/Total SNPs	%	Binomial test *p*-value
$p \leq 0.001$	180/316	57.0	0.0077	175/316	55.4	0.032	179/316	56.6	0.010
$0.001 < p \leq 0.01$	877/1624	54.0	0.00068	861/1624	53.0	0.0080	886/1624	54.6	0.00013
$0.01 < p \leq 0.5$	2556/5053	50.6	0.21	2604/5053	51.5	0.015	2588/5053	51.2	0.043
$0.5 < p \leq 1$	2535/5039	50.3	0.34	2532/5039	50.2	0.37	2519/5039	50.0	0.51

Table 2

Novel T2D susceptibility loci achieving genome-wide significance ($p<5\times10^{-8}$), identified through trans-ethnic "discovery" 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" meta-analysis of an additional 21,491 cases and 55,647 controls of European ancestry, genotyped with the Metabochip.

Locus	Lead SNP	Chr	Build 36 position (bp)	Alleles[a]		Trans-ethnic "discovery" meta-analysis		European ancestry "validation" meta-analysis		Combined meta-analysis	
				Risk	Other	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
TMEM154	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}
SSR1/RREB1	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}
FAF1	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}
POU5F1/TCF19	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}
LPP	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}
ARL15	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}
MPHOSPH9	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}

Chr: chromosome. OR: odds-ratio. CI: confidence interval.

[a] Alleles are aligned to the forward strand of NCBI Build 36.

Table 3

Properties of the 99% credible set of SNPs at ten established T2D susceptibility loci on the basis of association summary statistics from: (i) the meta-analysis of European ancestry GWAS only (12,171 cases and 56,862 controls); and (ii) 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).

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

Chr: chromosome. SNPs: number of SNPs.