doi:10.1038/nature18642

Supplementary Table 1a | Whole genome sequence cohort information

Ancestry	Study	Citation(s)	T2D Case Ascertainment	T2D Control Ascertainment	T1D and MODY exclusion criteria	Genotyping array
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Valle, T. et al. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. Diabetes Care 21(6), 949-958 (1998); Scott, L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316(5829), 1341-1345 (2007)	- Unrelated cases selected from FUSION families and stage 2 replication - Samples met 1999 World Health Organization (WHO) criteria of fasting plasma glucose ≥7.0 mmol/l or postload glucose during an OGTT ≥11.1 mmol/l, by report of diabetes medication use, or based on medical record review - Prioritized FUSION families with ≥2 first-degree relatives with T2D; BMI ≥18.5kg/m^2; case with GWAS data or earliest age at onset, if no GWAS data available - Prioritized FUSION stage 2 replication set with Metabochip data; BMI ≥18.5kg/m^2; earliest age of onset; age of onset ≥35	- Unrelated controls with normal glucose tolerance (NGT) based on WHO (1999) definitions: fasting plasma glucose <6.1 mM and 2 hour postload glucose during an OGTT <7.8 mM - Frequency matched to cases by birth province; BMI ≥18.5kg/m^2; age ≤80 - Within each birth province, prioritized samples from stage 2 replication with highest values for age + 2*BMI	- When possible, we prioritized cases with age of diagnosis between 35 and 60, without history of insulin-dependent diabetes among first degree relatives, with at least one full sibling diagnosed with T2D, and with at least one parent who was apparently nondiabetic.	Illumina 317K, HumanOmniExpress-12v1, and HumanExome-12v1_A
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	Wichmann, H. E., Gieger, C. and Illig, T. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67 Suppl 1, 26–30 (2005) Holle R, Happich M, Löwel H, Wichmann HE, MONICA/KORA Study Group. KORA - a research platform for population based health research. Gesundheitswesen. 2005 Aug; 67 Suppl 1:S19-25. PMID: 16032513	- Samples drawn from KORA F3 and F4 - Diabetic status validated by doctor or by medication use Prioritized cases with ≥1 first-degree relative with T2D (self-reported).- Cases have ≥1 first degree relative with type 2 diabetes (self-reported) - Cases have either BMI ≤30 and age of onset <65, or BMI ≤33 and age of onset ≤60	- Controls selected from KORA F4 - All controls are normal glucose tolerant: fasting glucose level <6.1 mmol/l and two hour glucose level after oral glucose tolerance test <7.8 mmol/l - Controls are either >60 years of age with BMI >32 or over 65 years of age with BMI >31	- None applied	Illumina HumanOmniExpress-12v1, Illumina HumanOmni2.5-8, Affymetrix Axiom array
European	UK Type 2 Diabetes Genetics Consortium (UKT2D)	Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–78 (2007); Voight, B.F. et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. 42, 579–589 (2010); Spector, T.D. and Williams, F.M. The UK Adult Twin Registry (TwinsUK). Twin Res. Hum. Genet. 9, 899–906 (2006)	- Cases drawn from the Wellcome Trust Case Control Consortium (WTCCC) - Female samples with age of diagnosis ≥66 years or BMI ≥32kg/m^2 excluded; male samples with age of diagnosis ≥62 years or BMI ≥31kg/m^2 excluded - Remaining samples were ranked by age and BMI, and the two ranks multiplied.	- Unrelated samples selected as controls from the Twins UK study - A twin pair was considered for selection if there was no recorded family history of diabetes, neither twin was ever recorded as impaired glucose tolerant (defined as fasting glucose >6.1mmol/L in any reading), there were available quantitative trait and genetic (GWAs) data, and no evidence of admixture in MDS analysis of GWAs data - From set of qualifying twin pairs, the best control twin was selected from each pair with the lowest ratio of fasting glucose level to BMI across all readings, and further prioritization of the qualifying unrelated samples involved selecting samples that had the lowest fasting glucose to (BMI * age) ratios - Top two principal components were used to perform pairwise sample matching between cases and possible controls, and the best control for each case was selected	- Cases having a first degree relative with type 1 diabetes; testing positive for GAD antibodies; or known to have other forms of diabetes, such as MODY, excluded - Controls having a first degree relative with type 1 diabetes excluded	Affymetrix GeneChip Human Mapping 500K Array Set (cases) and Illumina 317K (controls)
European	Malmo-Botnia Study	Guey LT et al. Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. Genet Epidemiol 35, 236–46 (2011); Groop, L. et al. Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. Diabetes 45, 1585–93 (1996); Parker, A. et al. A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. Diabetes 50, 675–80 (2001); Lyssenko, V. et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. NEJM 359, 2220–32 (2008); Lindholm, E., Agardh, E., Tuomi, T., Groop, L. & Agardh, C. D. Classifying diabetes according to the new WHO clinical stages. Eur. J. of Epid. 17, 983–9 (2001); Berglund, G. et al. Long-term outcome of the Malmo Preventive Project: Mortality and cardiovascular morbidity. J. of Intern. Med. 247, 19–29 (2000)	- A liability score was generated (Guey LT et al. 2011) which measures risk to T2D in the context of three known risk factors (age at onset, BMI, and gender) in 27,500 individuals drawn from three prospective cohorts: the Malmö Preventive Project, the Scania Diabetes Registry, and the Botnia Study; only BMI and gender used to construct scores for Scania and Botnia studies - Early-onset cases with low BMI and older controls with high BMI were prioritized	- Controls selected from the extreme of a liability score distribution, based upon gender, age and BMI at last follow-up visit; only BMI and gender used to construct scores for Malmö study - To match for ethnicity, equal numbers of controls were selected from the Botnia and Malmö studies	- Diabetic individuals with age of onset <35 years excluded	Illumina HumanOmniExpress-12v1

Supplementary Table 1b | Whole genome sequence sample characteristics

			Cases				Controls			
Ancestry	**Study**	**N Total**	**N Case**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**	**N Control**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	979	493	41.5	57.6 (7.9)	30.9 (5.6)	486	45.2	63.0 (7.2)	28.0 (3.9)
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	205	101	44.5	61.4 (8.2)	28.2 (2.8)	104	66.3	69.6 (5.6)	34.4 (3.5)
European	Malmo-Botnia Study	829	410	51.5	53.8 (9.7)	24.2 (2.6)	419	44.1	67.2 (7.7)	32.8 (4.0)
European	UK Type 2 Diabetes Genetics Consortium (UKT2D)	644	322	46.2	50.1 (8.4)	26.6 (2.7)	322	82.2	60.6 (10.0)	30.5 (5.8)

Supplementary Table 2 | Single-variant T2D association analysis descriptions

Dataset	Statistical Test		Covariates	Imputation
	Sample-size meta	Inverse variance meta		Quality filter
GoT2D sequenced	Score	Firth	Age, sex, before/after association*, PC1, PC2	N/A
GoT2D imputed				
DGDG	Score	Score	Age, gender, PC1, PC2	INFO>0.4
DGI	Score	Firth	Age, sex, BMI, 3 indicators (Malmo, Helsinki, Skara)	INFO>0.4
EGCUT 370K	LRT	LRT	Age, sex, PC1 – PC10	INFO>0.4
EGCUT OMNI	LRT	LRT	Age, sex, PC1 – PC10	INFO>0.4
FHS (MAF>0.01 only)	GEE (Wald)	GEE (Wald)	AGE, SEX, cohort, PC1 – PC10	INFO>0.4
FUSION	Score	Firth	Age (continuous), sex, 12 birth provinces	RSQ>0.3
INTERACT	Score	Score	Sex, PC1 – PC10	INFO>0.4
KORA	Score	Score	Age, sex, PC1 – PC3	INFO>0.4
MSSMIPM (Affymetrix)	Score (MAC≥200) Firth (MAC<200)	Score (MAC≥200) Firth (MAC<200)	Age, sex, PC1 – PC5	INFO>0.4
MSSMIPM (Illumina)	Score (MAC≥200) Firth (MAC<200)	Score (MAC≥200) Firth (MAC<200)	Age, sex, PC1 – PC4	INFO>0.4
PIVUS	Score	Score	Age, sex, PC1, PC2	INFO>0.4
ULSAM	Score	Score	PC1, PC2 (all males, same age)	INFO>0.4
WTCCC	Score	Score		INFO>0.4

LRT Likelihood ratio test, GEE Generalized estimating equations.
*Indicator function to account for observed temporal stratification based on sequencing date and center.

Supplementary Table 3a | Imputed cohort information

Ancestry	Study	Citation(s)	T2D Case Ascertainment	T2D Control Ascertainment	T1D and MODY exclusion criteria	Genotyping array
European	MT. SINAI BioMe Biobank Platform (BioMe (Affy))	Gottesman O, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present and future. Genet. Med. 15,761-771 (2013)	- From longitudinal EMR: random glucose ≥ 200 mg/dl ever, or physician-entered diagnosis (≥2 occurrences on 2 separate days), or T2D medication (≥2 occurrences on 2 separate days) - Age ≥ 25 years	- Age ≥25 years - Not having T2D - All available fasting glucose measurements <100 mg/dl	- Age ≥ 25 years	Affymetrix 6.0
European	MT. SINAI BioMe Biobank Platform (BioMe (Illumina))	Gottesman O, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present and future. Genet. Med. 15,761-771 (2013)	- From longitudinal EMR: random glucose ≥ 200 mg/dl ever, or physician-entered diagnosis (≥2 occurrences on 2 separate days), or T2D medication (≥2 occurrences on 2 separate days) - Age ≥ 25 years	- Age ≥25 years - Not having T2D - All available fasting glucose measurements <100 mg/dl	- Age ≥ 25 years	Illumina OMNI ExpressExome
European	Diabetes Gene Discovery Group (DGDG)	Sladek, R. et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445, 881-885 (2007)	- 1997 American Diabetes Association (ADA) criteria - Family history of diabetes in first-degree relatives - BMI<30 kg/m²	- Age at examination >45 years - Normal fasting glucose according to 1997 ADA criteria: FG <5.7 mmol/l - BMI <27 kg/m²	- Cases with age of diagnosis <45 years were screened for known MODY mutations - Cases from the Corbeil-Essonnes Hospital tested for fasting C-peptide levels; if fasting C-peptide<0.4 mg/l, subjects tested for anti-GAD antibodies; those with anti-GAD antibodies >10U/ml excluded. - 29% of cases from UMR8090 CNRS tested for anti-islet-antibodies and/or anti-insulin antibodies; positive cases excluded	Illumina Human Hap 300 Bead Array
European	Diabetes Genetics Initiative (DGI)	Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research , et al. Genome-Wide Association Analysis Identifies Loci for Type 2 Diabetes and Triglyceride Levels. Science 316,1331-1336 (2007)	- WHO (1999) criteria with fasting glucose ≥7.0 mmol/l or 2-hour glucose ≥11.1 mmol/l during an oral glucose tolerance test. - Age of diagnosis >35 years	- No first-degree relatives with T2D	- Anti-GAD antibodies <32 IU/ml in the Finnish samples and <1.3 anti-GAD relative units in the Swedish samples - Age of diagnosis >35 years	Affymetrix GeneChip Human Mapping 500k Array Set
European	Estonian Genome Center, University of Tartu (EGCUT-OMNI)	Leitsalu L, et al. Cohort profile: Estonian Biobank of the Estonian genome Center, University of Tartu. Int. J. Epidemiol. 44, 1137-1147 (2014)	- Standardized health examination together with questionnaires on health-related topics as described in WHO ICD-10. Data are regularly updated through linkage to national databases and registries.	- Random subset of the Estonian population	-None applied	Illumina OmniExpress Array
European	Estonian Genome Center, University of Tartu (EGCUT-370)	Leitsalu L, et al. Cohort profile: Estonian Biobank of the Estonian genome Center, University of Tartu. Int. J. Epidemiol. 44, 1137-1147 (2014)	- Standardized health examination together with questionnaires on health-related topics as described in WHO ICD-10. Data are regularly updated through linkage to national databases and registries.	- Random subset of the Estonian population	-None applied	Illumina HumanHap 370K Array
European	Framingham Heart Study (FHS)	Dawber, TR et al. Epidemiological approaches to heart disease: the Framingham Study. Am. J. Public Health Nations Health 41, 279-281 (1951); Feinleib, M et al. The Framingham Offspring Study. Design and preliminary data. Prev. Med. 4, 518-525 (1975); Splansky, GL et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am. J. Epidemiol.165,1328-1335 (2007)	- Gen 1 cohort: casual glucose ≥200 mg/dl or taking diabetes medication at any examination - Gen 2 cohort (offspring): FG≥126 mg/dl or diabetes treatment at any examination - Gen 3 cohort: fasting glucose ≥126 mg/dl or diabetes treatment at examination	- Fasting glucose <126 mg/dl and no T2D medication at the most recent study visit	-None applied	Affymetrix GeneChip Human Mapping 500k Array Set + MIPS 50K
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Scott, L. J. et al. (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316, 1341-1345 (2007)	- WHO 1999 criteria of FG≥7.0 mmol/l or 2-hour plasma glucose ≥11.1 mmol/l or reported diabetes medication use or based on medical record review	- NGT as defined by WHO 1999 criteria	- No known or probable type 1 diabetes among first degree relatives '-Insulin treatment initiated within 10 years of disease diagnosis, detectable levels of anti-GAD antibodies and fasting C-peptide ≤0.30 nmol/l '- Insulin treatment initiated within 4 years of diagnosis and fasting C-peptide ≤0.30 nmol/l	Illumina Human Hap 300 Bead Array
European	InterAct	Langenberg, C Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. Diabetlogia 54, 2272-2282 (2011)	- Self-report (self-reported history of T2D, doctor-diagnosed T2D, diabetes drug use) and linkage to primary care registers, secondary care registers, medication use (drug registers), hospital admissions, and mortality data, using information from any follow-up visit or external evidence with a date later than the baseline visit. -In Sweden, cases ascertained via local and national diabetes and pharmaceutical registers. -Centers outside Sweden required evidence of T2D from ≥2 independent sources, including individual medical records review at some centers.	- Individuals without T2D	-None applied	Illumina HumanHap 660 Array
European	KORAgen Study Helmholtz zentrum München (KORA)	Wichmann HE, Gieger C, Illig T MONICA/KORA Study Group. Gesundheitswesen 67 Suppl 1, S26-30. Review (2005)	- Self-report in personal interview validated by a questionnaire mailed to the treating physician and/or by medical chart review -	- Non-diabetic by self-report	-None applied	Affymetrix GeneChip Human Mapping 500k Array Set
European	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	Lind, L. et al. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. Arterioscler. Thromb. Vasc. Biol. 25, 2368-2375 (2005)	- Fasting blood glucose >6.1 mmol/l or known diabetes	- Individuals without T2D	-None applied	Illumina Metabochip and Illumina OmniExpress Array.
European	Uppsala Longitudinal Study of Adult Men (ULSAM)	Ingelsson E, et al. Insulin resistance and risk of congestive heart failure. JAMA 294, 334-341 (2005)	- Hospital discharge register-defined diabetes before 2002.	- Individuals without T2D	-None applied	Illumina Metabochip and Illumina OmniExpress Array.
European	Welcome Trust Case Control Consortium (WTCCC)	Zeggini, E. et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat. Genet. 40, 638-645 (2008)	- Current prescribed treatment with oral agents and/or insulin, or, for individuals treated with diet alone, laboratory evidence of hyperglycemia as defined by WHO - All cases were diagnosed between ages 25 and 75 years.	- Selected without reference to T2D status - 1958 Birth Cohort controls of self-reported white ethnicity and representative of gender and each geographical region - UK blood donor controls selected based on sex and geographical region to reproduce the distribution of the samples of the 1958 Birth Cohort	- Age of diagnosis ≥25 - Absence of first-degree relatives with T1D - Individuals with other known forms of diabetes were excluded - ≥ 1 year between diagnosis and institution of regular insulin therapy - Negative testing for anti-GAD antibodies	Affymetrix Human Mapping 500K Array

Supplementary Table 3b | Imputed cohort sample characteristics

			Cases					Controls			
Ancestry	**Study**	**N Total**	**N Case**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**	**Mean age of diagnosis (SD), years**	**N Control**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**
European	Diabetes Gene Discovery Group (DGDG)	1374	677	266 (39.3)	59.5 (10.0)	26.1 (2.7)	45.4(8.3)	697	416 (59.7)	53.5 (5.7)	23.2 (1.8)
European	Diabetes Genetics Initiative (DGI)	1956	899	419 (46.6)	65.3 (9.9)	28.1 (4.1)	59.2 (10.1)	1057	533 (49.6)	58.3 (9.6)	26.7 (3.7)
European	Estonian Genome Center, University of Tartu (EGCUT-370)	1848	80	39 (48.8)	61.9 (11.3)	31.6 (4.8)	NA	1768	902 (51.0)	39.7 (16.1)	25.7 (5.1)
European	Estonian Genome Center, University of Tartu (EGCUT-OMNI)	6402	389	228 (58.6)	70.2 (11.8)	31.1 (6.2)	NA	6013	3259 (54.2)	50.9 (20.4)	26.5 (5.1)
European	Framingham Heart Study (FHS)	8333	673	287 (42.6)	63.7 (12.4)	31.4 (6.5)	NA	7660	4219 (55.1)	52.3 (16.0)	27.0 (5.1)
European	Finland United States Investigation of NIDDM (FUSION) Study	2150	1060	457 (43.1)	63.0 (7.6)	30.3 (4.7)	NA	1090	560 (51.3)	63.2 (7.4)	26.9 (3.7)
European	INTERACT	9292	4624	2395 (51.8)	NA	30.0 (4.8)	NA	4668	2995 (64.2)	NA	25.9 (4.2)
European	KORAgen Study Helmholtz zentrum München (KORA)	3978	993	447(45.1)	60.4 (11.0)	31.0 (5.4)	NA	2985	1558 (52.2)	55.6 (13.2)	27.3 (4.5)
European	MT. SINAI BioMe Biobank Platform (BioMe (Affy))	587	132	35 (26.5)	66.0 (10.7)	31.3 (6.3)	NA	455	158 (34.7)	66.0(10.7)	26.5 (5.2)
European	MT. SINAI BioMe Biobank Platform (BioMe (Illumina))	1902	255	74 (29.0)	69.1 (9.0)	30.7 (6.2)	NA	1647	846 (51.4)	69.1 (9.0)	26.3 (5.0)
European	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	949	111	46 (41.4)	70.2 (0.1)	29.1 (5.3)	NA	838	492 (51.2)	70.2 (0.2)	26.8 (4.2)
European	Uppsala Longitudinal Study of Adult Men (ULSAM)	1119	166	0	71.0 (0.7)	27.9 (3.9)	NA	953	0	71.0(0.6)	26.0(3.2)
European	Welcome Trust Case Control Consortium (WTCCC)	4524	1586	649 (40.9)	58.3 (10.1)	32.2 (6.2)	49.0 (11.9)	2938	1492 (50.8)	43.3 (12.3)	28.2 (4.3)

Supplementary Table 4a | Exome sequence cohort information

Ancestry	Study	Citation(s)	T2D Case Ascertainment	T2D Control Ascertainment	T1D and MODY exclusion criteria	Genotyping array
African American	Jackson Heart Study (AJ)	Taylor, H. A. et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn Dis 15, S6–4 (2005)	- Genomic DNA from blood and genome-wide genotypes available (minimum 25 ug DNA remaining) - Non-related individuals based on family IDs - ADA 2004 definition of T2D diagnosed either of two exams	- No T2D by ADA 2004 definition, fasting plasma glucose <100 mg/dl, and HbA1c <6% at each of two exams - Controls were matched to cases in a two-stage approach: 1. Strong matches (greedy algorithm): age > 50, sex match, BMI within 1 unit, and age within 5 years (N=457 matched pairs) 2. Closest available matches: sex match and BMI > 25; for females, BMI within 5 units and age within 20 years; for males, BMI within 8 units and age within 25 years (N=117 matched pairs)	- Samples with age of onset <16 years and treated only with insulin were excluded	Affymetrix Genome-Wide Human SNP Array 6.0
African American	Wake Forest School of Medicine Study (AW)	Palmer, N. D. et al. A genome-wide association search for type 2 diabetes genes in African Americans. PLoS One 7:e29202. (2012)	- Cases are self-reported diabetics with diabetic nephropathy, recruited from dialysis clinics - Age of onset ≥25 - Individuals excluded if at any point after diagnosis treatment consisted of insulin therapy alone - Additionally, at least one of the following three criteria met for inclusion: a) T2D diagnosed at least 5 years before initiating renal replacement therapy, b) background or greater diabetic retinopathy, and c) ≥100 mg/dl proteinuria on urinalysis in the absence of other causes of nephropathy	- No current diagnosis of diabetes or renal disease - Individuals recruited from community and internal medicine clinics	- Type 1 diabetes is uncommon in African Americans, so no ICA tests have been performed - Age of onset ≥25 - Individuals excluded if at any point after diagnosis treatment consisted of insulin therapy alone	Affymetrix Genome-Wide Human SNP Array 6.0
East Asian	Korea Association Research Project [Korean] (EK)	Cho, Y. S. et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat. Genet. 41, 527–534 (2009)	- Past history of T2D - Use of T2D medication - Fasting plasma glucose ≥7 mmol/l or plasma glucose ≥11.1 mmol/l 2 hours after ingestion of 75gm oral glucose load - Age of disease onset ≥40 years - Participants with early onset and family history prioritized	- No past history of diabetes - No anti-diabetic medication - Fasting plasma glucose <5.6 mmol/l and plasma glucose 2 hours after ingestion of 75g oral glucose load <7.8 mmol/l at both baseline and follow up timepoints - Older subjects with normal glucose prioritized	- Samples with age of diagnosis <40 excluded	Affymetrix Genome-Wide Human SNP Array 5.0
East Asian	Singapore Diabetes Cohort Study and Singapore Prospective Study Program [Singapore Chinese] (ES)	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)	- Clinically ascertained T2D from primary care clinics - Individuals with early age of diagnosis and with at least one first degree relative with T2D were preferentially selected	- Fasting blood glucose <6 mmol/l - No personal history of diabetes - No anti-diabetic medication - Older controls preferentially selected	- Clinical records were extracted from primary care clinics, and suspected T1D and MODY cases were excluded	Illumina Human610-Quad BeadChip / Illumina Human1M-Duo v3.0
European	Ashkenazi (UA)	Atzmon, G. et al. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. PLoS Biol. 4(4), e113 (2006); Atzmon, G. et al. Evolution in health and medicine Sackler colloquium: Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. Proc Natl Acad Sci U S A. 107 (Suppl 1), 1710-1717 (2010); Permutt, M.A. et al. A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. Diabetes 50(3), 681-685 (2001); Blech et al. Predicting diabetic nephropathy using a multifactorial genetic model. PLoS One 6(4), e18743 (2011)	- Ashkenazi Jewish origin, defined as having all four grandparents born in Northern or Eastern Europe; subjects with known or supected Sephardic Jewish or non-Jewish ancestry excluded - T2D defined according to the World Health Organisation criteria (fasting glucose >140 mg/dl on two or more ocasions or random glucose >200mg/dl) - To avoid late-onset T1D, patients who became insulin-dependent within 2 years of diagnosis excluded; anti-GAD or anti-islet cell antibody titers not routinely measured - T2D cases were selected from two separate DNA collections: 1. Genome-wide, affected-sibling-pair linkage study (Permutt et al. Diabetes 2001). Families in which both parents were known to have diabetes were excluded. One affected individual selected from each family and, wherever possible, sibling with youngest age of diagnosis selected. 2. Study to determine genetic risk for diabetic complications (Blech et al. PLoS One 2011). Patients ascertained by the Israel Diabetes Research Group between 2002 and 2004 from 15 diabetes clinics throughout Israel. Primary selection criteria: (1) known T1D or T2D for more than 10 years, (2) 4 grandparents being either Ashkenazi or Sephardic-North African Jewish. For this study, only T2D patients with all 4 grandparents of Ashkenazi Jewish origin and age of diagnosis between 35 and 60 were selected.	- Fasting blood glucose <7 mmol/l - No personal history of diabetes - No anti-diabetic medications	- Patients who became insulin-dependent within 2 years of diagnosis excluded	Illumina Cardio-Metabo Chip
European	Metabolic Syndrome in Men Study [Finnish] (UM)	Stancakova, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes 58, 1212–1221 (2009)	- Previous diagnosis of T2D, or both fasting and 2-hr criteria met for new T2D diagnosis - C-peptide >0.10 nmol/L - Anti-GAD antibody <50 U/mL to rule out T1D - Family history of diabetes (parents, sibs, children, grandparents, avuncular, cousins) - Unrelated individuals based on family ID and IBS analyses - Preferentially select individuals with with genotype data (N=494), as well as non-genotyped individuals with earlier possible age of diagnosis (N=26)	- Normal glucose tolerance at baseline and follow-up visits - Prioritized samples with no family history of diabetes and meeting strict NGT criteria: fasting glucose <5.6 mmol/l and 2 hour post-challenge glucose <7.8 mmol/l - Additional samples selected with fasting glucose <6.1 mmol/l and 2 hour post-challenge glucose <7.8 mmol/l - Unrelated samples - Older controls preferentially selected	- C-peptide <0.10 nmol/L - Anti-GAD antibody >50 U/mL	Illumina Cardio-Metabo Chip and HumanOmniExpress-12v1
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study [Finnish]	Valle, T. et al. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. Diabetes Care 21(6), 949-958 (1998); Scott, L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316(5829), 1341-1345 (2007)	- Unrelated cases selected from FUSION families and stage 2 replication - Samples met 1999 World Health Organization (WHO) criteria of fasting plasma glucose ≥7.0 mmol/l or postload glucose during an OGTT ≥11.1 mmol/l, by report of diabetes medication use, or based on medical record review - Prioritized FUSION families with ≥2 first-degree relatives with T2D; BMI ≥18.5kg/m^2; case with GWAS data or earliest age at onset, if no GWAS data available - Prioritized FUSION stage 2 replication set with Metabochip data; BMI ≥18.5kg/m^2; earliest age of onset; age of onset ≥35	- Unrelated controls with normal glucose tolerance (NGT) based on WHO (1999) definitions: fasting plasma glucose <6.1 mM and 2 hour postload glucose during an OGTT <7.8 mM - Frequency matched to cases by birth province; BMI ≥18.5kg/m^2; age ≤80 - Within each birth province, prioritized samples from stage 2 replication with highest values for age + 2*BMI	- When possible, we prioritized cases with age of diagnosis between 35 and 60, without history of insulin-dependent diabetes among first degree relatives, with at least one full sibling diagnosed with T2D, and with at least one parent who was apparently nondiabetic.	Illumina 317K, HumanOmniExpress-12v1, and HumanExome-12v1_A
European	KORA [German]	Wichmann, H. E., Gieger, C. and Illig, T. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67 Suppl 1, 26–30 (2005) Holle R, Happich M, Löwel H, Wichmann HE, MONICA/KORA Study Group. KORA - a research platform for population based health research. Gesundheitswesen. 2005 Aug; 67 Suppl 1:S19-25. PMID: 16032513	- Samples drawn from KORA F3 and F4 - Diabetic status validated by doctor or by medication use - Cases have ≥1 first degree relative with type 2 diabetes (self-reported) - Cases have either BMI ≤30 and age of onset <65, or BMI ≤33 and age of onset ≤60	- Controls selected from KORA F4 - All controls are normal glucose tolerant: fasting glucose level <6.1 mmol/l and two hour glucose level after oral glucose tolerance test <7.8 mmol/l - Controls are either >60 years of age with BMI >32 or over 65 years of age with BMI >31	- None applied	Illumina HumanOmniExpress-12v1, Illumina HumanOmni2.5-8, Affymetrix Axiom array
European	UKT2D Consortium	Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–78 (2007); Voight, B.F. et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. 42, 579–589 (2010); Spector, T.D. and Williams, F.M. The UK Adult Twin Registry (TwinsUK). Twin Res. Hum. Genet. 9, 899–906 (2006)	- Cases drawn from the Wellcome Trust Case Control Consortium (WTCCC) - Female samples with age of diagnosis ≥66 years or BMI ≥32kg/m^2 excluded; male samples with age of diagnosis ≥62 years or BMI ≥31kg/m^2 excluded - Remaining samples were ranked by age and BMI, and the two ranks multiplied. 356 samples with the lowest values for this rank multiplier were selected for initial inclusion in the study	- Unrelated samples selected as controls from the Twins UK study - A twin pair was considered for selection if there was no recorded family history of diabetes, neither twin was ever recorded as impaired glucose tolerant (defined as fasting glucose >6.1mmol/L in any reading), there were available quantitative trait and genetic (GWAs) data, and no evidence of admixture in MDS analysis of GWAs data - From set of qualifying twin pairs, the best control twin was selected from each pair with the lowest ratio of fasting glucose level to BMI across all readings, and further prioritization of the qualifying unrelated samples involved selecting samples that had the lowest fasting glucose to (BMI * age) ratios - Top two principal components were used to perform pairwise sample matching between cases and possible controls, and the best control for each case was selected	- Cases having a first degree relative with type 1 diabetes; testing positive for GAD antibodies; or known to have other forms of diabetes, such as MODY, excluded - Controls having a first degree relative with type 1 diabetes excluded	Affymetrix GeneChip Human Mapping 500K Array Set (cases) and Illumina 317K (controls)

European	Malmo-Botnia Study [Finnish, Swedish]	Guey LT et al. Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. Genet Epidemiol 35, 236–46 (2011); Groop, L. et al. Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. Diabetes 45, 1585–93 (1996); Parker, A. et al. A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. Diabetes 50, 675–80 (2001); Lyssenko, V. et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. NEJM 359, 2220–32 (2008); Lindholm, E., Agardh, E., Tuomi, T., Groop, L. & Agardh, C. D. Classifying diabetes according to the new WHO clinical stages. Eur. J. of Epid. 17, 983–9 (2001); Berglund, G. et al. Long-term outcome of the Malmo Preventive Project: Mortality and cardiovascular morbidity. J. of Intern. Med. 247, 19–29 (2000)	- A liability score was generated (Guey LT et al. 2011) which measures risk to T2D in the context of three known risk factors (age at onset, BMI, and gender) in 27,500 individuals drawn from three prospective cohorts: the Malmö Preventive Project, the Scania Diabetes Registry, and the Botnia Study; only BMI and gender used to construct scores for Scania and Botnia studies - Eligible cases limited to individuals between 35 and 60 years of age and with a BMI between 20 and 35 - To match for ethnicity, 250 Botnia cases with the most extreme liability scores were selected, while 125 cases were selected from each of the Scania and Malmö studies	- Controls selected from the extreme of a liability score distribution, based upon gender, age and BMI at last follow-up visit; only BMI and gender used to construct scores for Malmö study - Eligible controls limited to individuals above 35 years of age at follow-up and with a BMI between 20 and 40 - To match for ethnicity, equal numbers of controls were selected from the Botnia and Malmö studies	- Diabetic individuals with age of onset <35 years excluded	Illumina HumanOmniExpress-12v1
Hispanic	San Antonio Mexican American Family Studies: San Antonio Family Heart Study, San Antonio Family Diabetes/Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and Family Investigation of Nephropathy and Diabetes Study - San Antonio Component (HA)	Mitchell, B. D. et al. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. Circulation 94, 2159–2170 (1996); Hunt, K. J. et al. Genome-wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio Family Diabetes/Gallbladder Study. Diabetes 54, 2655–2662 (2005); Coletta, D. K. et al. Genome-wide linkage scan for genes influencing plasma triglyceride levels in the Veterans Administration Genetic Epidemiology Study. Diabetes 58, 279–284 (2009); Knowler, W. C. et al. The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. J. Diabetes Complicat. 19, 1–9 (2005)	- Unrelated and non-overlapping individuals/samples drawn from four separate family studies, San Antonio, TX - Cases met one or more of four criteria: 1. American Diabetes Association [ADA] criterion (2002) -- fasting plasma glucose ≥126 mg/dl 2. World Health Organization [WHO] criteria (1999) -- fasting plasma glucose ≥126 mg/dl or a 2 hour glucose tolerance test ≥200 mg/dl 3. Self-reported physician-diagnosed diabetes and self-reported current therapy with either oral antidiabetic agents or insulin 4. Had hemoglobin A1c (HbA1c) ≥7.0%	- No self-reported antidiabetic therapy at any visit, including oral agents or insulin prescribed as a result of physician-diagnosed diabetes AND one or more of the following: 1. Fasting glucose <126 mg/dl at each visit 2. If OGTT performed, 2 hour glucose must be <200 mg/dl 3. No history of diabetes and HbA1c <6.0%, or HbA1c 6.0-6.9% and fasting glucose <126 mg/dl Samples prioritized with strict NGT with no family history first, then NGT in two visits, followed by oldest age	- None applied	Illumina Cardio-Metabo Chip
Hispanic	Starr County, Texas (HS)	Hanis, C. L. et al. Diabetes among Mexican Americans in Starr County, Texas. Am. J. Epidemiol. 118, 659–672 (1983); Below JE, et al. Genome-wide association and meta-analysis in populations from Starr County, Texas and Mexico City identify type 2 diabetes susceptibility loci and enrichment for eQTLs in top signals. Diabetologia 54, 2047-2055 (2011)	- Diagnosis of diabetes according to National Diabetes Data Group (1979) guidelines drawn from several studies in Starr County: 1. Fasting glucose ≥ 140 mg/dl on more than 1 occasion 2. Self-reported physician-diagnosed diabetes and self reported therapy with either oral antidiabetic agents or insulin (either currently or for more than one month in the past) - In instances where cases were drawn from families, the individual with youngest age at onset was chosen	- Controls ascertained from epidemiologically represented sample of individuals in Starr County, TX - Individuals with known diagnosis of diabetes excluded - Impaired glucose tolerant and impaired fasting glucose controls retained due to the age difference between cases and controls (controls are younger on average) and to allow sufficient sample size	- Cases with age of onset <35 and BMI<30 excluded as potential T1D or MODY	Affymetrix Genome-Wide Human SNP Array 6.0
South Asian	London Life Sciences Population Study [UK Indian Asians] (SL)	Chambers, J.C. et al. Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat. Genet. 41, 1170-1172 (2009); Chambers, J.C. et al. Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. Diabetes 58, 2703-2708 (2009); van der Harst, P. et al. Seventy-five genetic loci influencing the human red blood cell. Nature 492, 369-375 (2012)	- Samples chosen from a population-based cohort study - Indian Asians living in West London, UK with all 4 grandparents born on the Indian subcontinent - Prevalent T2D defined as previous physician diagnosis of diabetes on treatment, with onset of diabetes after the age of 18 years and without insulin use in the first year after diagnosis; or fasting plasma glucose ≥7.0 mmol/L	- No previous history of diabetes - No anti-diabetic medication - Fasting plasma glucose <6.0 mmol/L	- Samples with age of onset <18 excluded	Illumina Human610-Quad BeadChip
South Asian	Singapore Indian Eye Study [Singapore Indians] (SS)	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)	- HbA1c ≥6.5% or personal history of diabetes with age at diagnosis available - Preferentially selected cases with at least one first degree relative with T2D	- HbA1c <6% - No personal history of diabetes - Not taking antidiabetes medication - Older controls preferentially selected	- None applied	Illumina Human610-Quad BeadChip

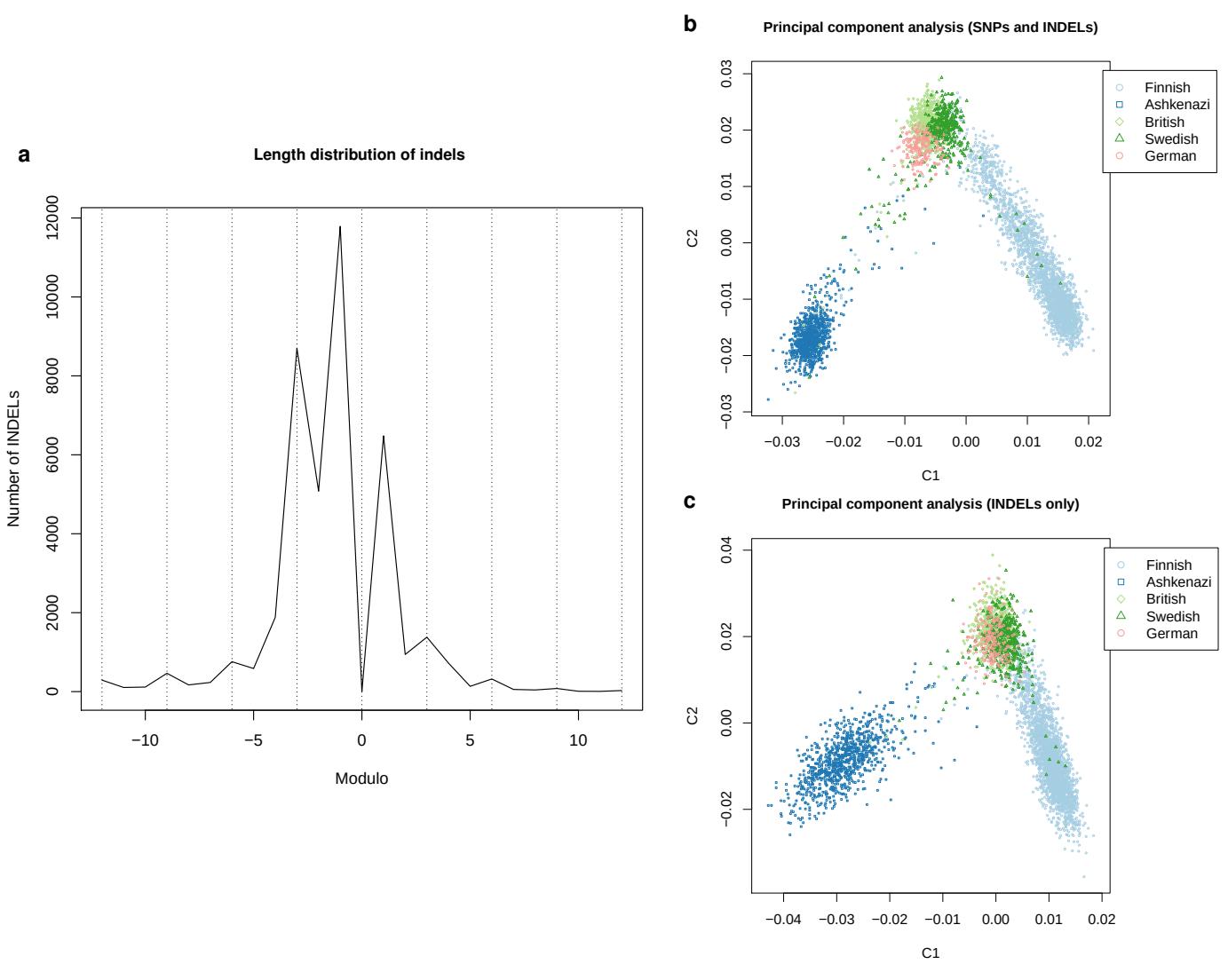
Supplementary Table 4b | Exome sequence sample characteristics

			Cases					Controls			
Ancestry	**Study**	**N Total**	**N Case**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**	**Mean age of diagnosis (SD), years**	**N Control**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**
African American	Jackson Heart Study (AJ)	1026	500	333 (66.6)	F: 58.1 (10.0) M: 57.9 (11.2)	F: 34.5 (6.8) M: 31 (4.7)	F: 49.1 (11.2) M: 49.3 (11.0)	526	333 (63.3)	F: 55.8 (11.4) M: 56.4 (11.2)	F: 33 (6.8) M: 30 (5.1)
African American	Wake Forest School of Medicine Study (AW)	1048	518	308 (59.5)	F: 64.8 (9.3) M: 62.3 (8.7)	F: 30.3 (7.1) M: 27.9 (5.3)	F: 46.9 (10.0) M: 47.6 (9.6)	530	297 (56.0)	F: 50.7 (12.2) M: 51.4 (10.8)	F: 31.5 (7.3) M: 28 (6.0)
East Asian	Korea Association Research Project [Korean] (EK)	1087	526	240 (45.6)	F: 55.4 (7.2) M: 52.5 (7.4)	F: 26.3 (3.4) M: 25.2 (3.0)	F: NA M: NA	561	328 (58.5)	F: 62.9 (3.5) M: 63.8 (3.6)	F: 24.2 (3.1) M: 23.1 (2.8)
East Asian	Singapore Diabetes Cohort Study and Singapore Prospective Study Program [Singapore Chinese] (ES)	1078	486	253 (52.1)	F: 58.7 (9.9) M: 57.3 (8.7)	F: 25.6 (3.9) M: 25.6 (3.7)	F: 45.5 (9.0) M: 44.2 (9.4)	592	363 (61.3)	F: 58.1 (6.7) M: 58.5 (7.5)	F: 22.8 (3.4) M: 23 (3.3)
European	Ashkenazi (UA)	861	506	238 (47)	F: 65.9 (8.7) M: 65.6 (8.6)	F: 27.6 (3.2) M: 27.2 (3.1)	F: 49 (5.5) M: 47.7 (5.3)	355	202 (56.9)	F: 80.1 (14.5) M: 76.7 (11.5)	F: 24.4 (4.3) M: 26.3 (3.7)
European	Metabolic Syndrome in Men Study [Finnish] (UM)	982	484	0 (0)	F: NA M: 60.4 (6.7)	F: NA M: 30.6 (5.1)	F: NA M: 54.6 (8.5)	498	0 (0)	F: NA M: 54.7 (4.5)	F: NA M: 25.8 (3.1)
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study [Finnish]	948	472	201 (42.6)	F: 58.6 (9.0) M: 56.9 (7.1)	F: 31.3 (5.6) M: 30.5 (5.5)	F: NA M: NA	476	214 (45.0)	F: 63.8 (7.1) M: 62.2 (7.3)	F: 28.5 (4.4) M: 27.5 (3.3)
European	KORA [German]	187	97	43 (44.3)	F: 60.6 (8.8) M: 61.9 (7.6)	F: 28.9 (2.8) M: 27.9 (2.8)	F: NA M: NA	90	57 (63.3)	F: 68.9 (5.4) M: 70.9 (5.8)	F: 34.6 (3.5) M: 34.2 (3.4)
European	UKT2D Consortium	642	322	147 (45.7)	F: 51.2 (9.1) M: 48.9 (7.7)	F: 27.1 (2.7) M: 26.4 (2.7)	F: NA M: NA	320	265 (82.8)	F: 60.9 (10.2) M: 59.8 (9.0)	F: 31 (6.2) M: 28.4 (3.7)
European	Malmo-Botnia Study [Finnish, Swedish]	921	478	262 (54.8)	F: 56.8 (10.1) M: 52.5 (10.2)	F: 24.8 (2.8) M: 23.9 (2.2)	F: 47.4 (7.7) M: 45.6 (7.2)	443	194 (43.8)	F: 68 (8.0) M: 65.6 (8.1)	F: 33.7 (4.1) M: 32.2 (3.9)
Hispanic	San Antonio Mexican American Family Studies: San Antonio Family Heart Study, San Antonio Family Diabetes/Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and Family Investigation of Nephropathy and Diabetes Study - San Antonio Component (HA)	490	272	160 (58.8)	F: 58 (13.3) M: 57.7 (11.0)	F: 32.9 (6.9) M: 31.6 (7.7)	F: 45.5 (14.7) M: 44.5 (14.4)	218	128 (58.7)	F: 53.3 (15.2) M: 50.9 (14.6)	F: 31.1 (7.3) M: 28.9 (6.1)
Hispanic	Starr County, Texas (HS)	1453	749	447 (59.7)	F: 56 (11.9) M: 56.9 (11.8)	F: 32.9 (6.8) M: 30.2 (5.3)	F: 46 (11.0) M: 47.7 (11.0)	704	506 (71.9)	F: 39.1 (9.4) M: 39.4 (11.1)	F: 30.4 (6.5) M: 29.5 (5.3)
South Asian	London Life Sciences Population Study [UK Indian Asians] (SL)	1069	531	75 (14.1)	F: 53.4 (5.5) M: 52.7 (5.7)	F: 27.7 (3.0) M: 26.5 (2.8)	F: NA M: NA	538	85 (15.8)	F: 63.6 (8.9) M: 63.4 (9.2)	F: 28.2 (4.4) M: 27 (3.3)
South Asian	Singapore Indian Eye Study [Singapore Indians] (SS)	1148	563	250 (44.4)	F: 59.8 (9.4) M: 61.6 (9.9)	F: 28.2 (6.0) M: 25.9 (4.0)	F: 50.5 (10.7) M: 50.9 (10.5)	585	288 (49.2)	F: 55.8 (9.7) M: 56.4 (10.4)	F: 26.3 (5.6) M: 24.4 (3.5)

Supplementary Figure 5 | Quality control of 12,940 WES samples. To assess the sequencing quality of each sample, we computed multiple statistics stratified by sample ethnicity. We then identified outlier samples relative to any of the statistical distributions and excluded them from further analysis. Shown are the distributions for nine representative statistics after samples were removed from analysis; note that these metrics are computed prior to any variant quality control and thus measure different statistics than presented in other display items. Number of variants: the number of variants (biallelic or multiallelic SNPs and INDELs) at which the sample exome carries a minor allele. Number of biallelic SNPs: the number of biallelic SNPs at which the sample exome carries a minor allele. Number of biallelic indels: the number of biallelic INDELs at which the sample exome carries a minor allele. Number of singletons: the number of variants carried by the sample alone (e.g., at which all other samples have the reference genotype). Heterozygosity: the heterozygosity of the sample computed across all variant sites. Heterozygosity at low frequency variants: the heterozygosity of the sample computed across low-frequency (MAF < 1%) variant sites. Heterozygous:Homozygous ratio: the ratio of heterozygous non-reference alleles to homozygous non-reference alleles in the sample. Mean allele balance: the fraction of sequence reads containing the non-reference allele, averaged over all heterozygous sites in the sample. Mean allele balance at singleton sites: the fraction of sequence reads containing the non-reference allele, averaged over all singleton heterozygous sites in the sample.

Number of variants
Number of SNPs
Number of Indels
Number of singletons
Heterozygosity
Heterozygosity at low frequency variants
Heterozygous.Homozygous.ratio
Mean allele balance
Mean allele balance at singleton sites
Sample Values
African−American
Korean
Chinese
Indian
Hispanic
Ashkenazi
Finnish
Swedish
German
British

Supplementary Figure 6 | **Quality control of INDELs.** To assess the quality of called INDEL variants, we computed two metrics. (a) The number of INDELs with size equal to x (mod 3), for various values of x. Negative values represent deletions, while positive values represent insertions. As frameshift variants are more likely to disrupt protein sequence than in-frame deletions, spikes at increments of three are expected for INDEL variants in the population. (bc) Principal component analysis of the 12,940 European samples, computed using common (MAF > 1%) (b) SNPs and INDELs and (c) INDELs only. If the majority of common SNPs and INDELs are of high quality, the principal components should be concordant between the two analyses.

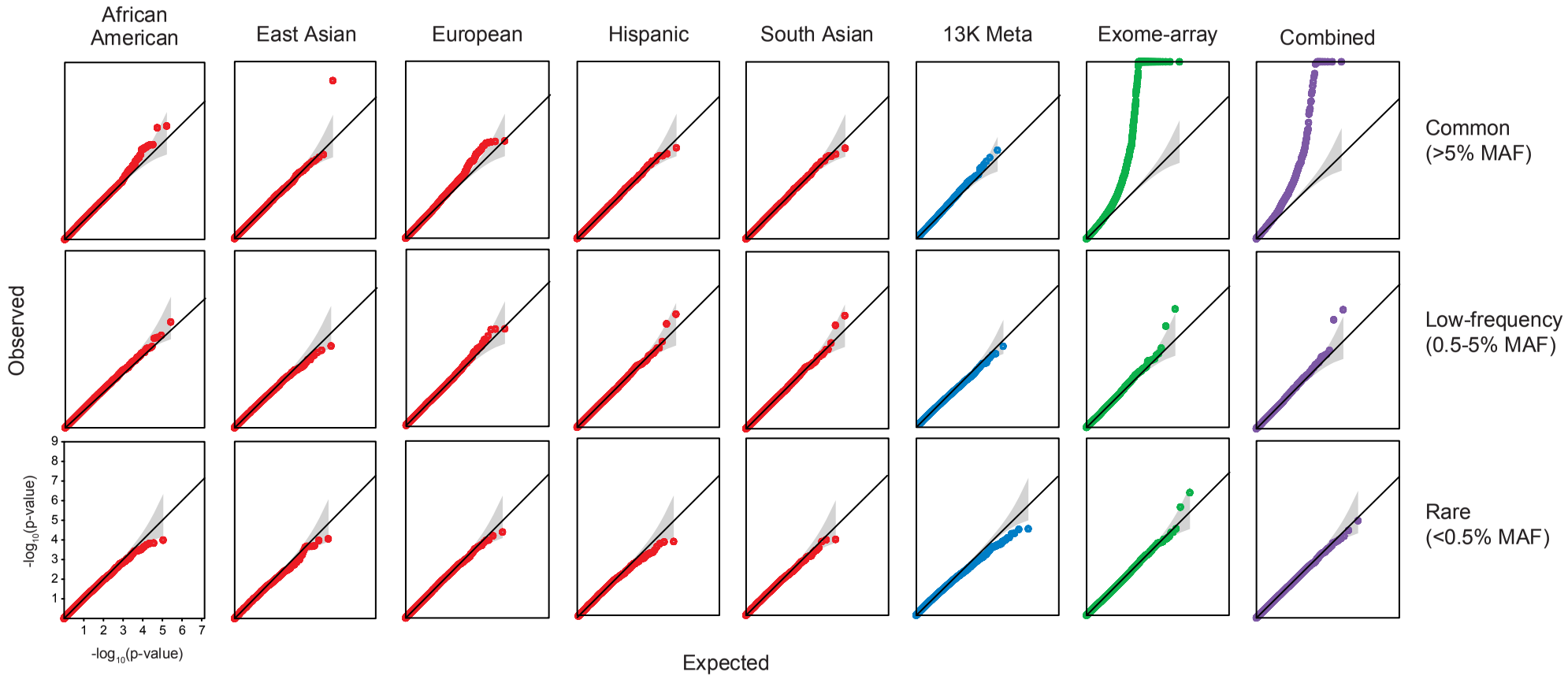

a
Length distribution of indels
Number of INDELs
0
2000
4000
6000
8000
10000
12000
−10
−5
0
5
10
Modulo
b
Principal component analysis (SNPs and INDELs)
C2
C1
−0.03
−0.02
−0.01
0.00
0.01
0.02
0.03
Finnish
Ashkenazi
British
Swedish
German
c
Principal component analysis (INDELs only)
C2
C1
−0.04
−0.03
−0.02
−0.01
0.00
0.01
0.02
0.04
Finnish
Ashkenazi
British
Swedish
German

Supplementary Table 7 | Per individual variant counts by ethnic group

QC+ variants

	All samples	African-American	East-Asian	European	Hispanic	South-Asian
Samples:	12,940	2,074	2,165	4,541	1,943	2,217
Variants:						
Synonymous SNP	9243 [8423;11487]	10910 [9857;11487]	8837 [8559;9258]	8851 [8427;9264]	9024 [8595;9428]	9070 [8423;9387]
Missense SNP	7636 [6935;9271]	8885 [8038;9271]	7348 [7042;8133]	7341 [6939;7933]	7468 [7069;8155]	7499 [6935;7885]
Start SNP*	11 [4;22]	12 [5;22]	9 [4;16]	11 [4;18]	11 [5;18]	11 [5;20]
Nonsense SNP*	62 [37;93]	69 [50;93]	60 [40;82]	60 [37;84]	61 [45;93]	60 [44;81]
Frameshift INDEL*	137 [115;182]	147 [124;172]	137 [120;160]	134 [115;155]	136 [118;182]	136 [117;156]
Inframe INDEL	76 [56;111]	89 [70;111]	72 [60;90]	72 [56;91]	75 [58;92]	75 [59;91]
3'UTR SNP, INDEL	532 [449;700]	633 [548;700]	513 [465;589]	506 [449;567]	516 [461;578]	523 [474;576]
5'UTR SNP, INDEL	864 [747;1119]	1021 [903;1119]	836 [783;895]	824 [747;904]	843 [781;942]	848 [769;924]
Intron SNP, INDEL	13110 [11477;16462]	15590 [13917;16462]	12540 [12119;13972]	12510 [11989;13611]	12810 [12159;14405]	12830 [11477;13357]
Essential splicing SNP, INDEL*	40 [27;61]	46 [30;61]	42 [30;55]	38 [27;55]	40 [27;56]	39 [28;56]
Other splicing SNP, INDEL	1586 [1411;2022]	1886 [1675;2022]	1531 [1455;1652]	1514 [1411;1614]	1543 [1426;1641]	1547 [1417;1644]
Non-coding RNA SNP, INDEL	245 [192;337]	288 [241;337]	232 [192;266]	236 [196;290]	238 [199;277]	239 [193;283]
All variants	**34070 [30901;41971]**	**40330 [36180;41971]**	**32640 [31659;35250]**	**32570 [31079;35021]**	**33280 [31666;36390]**	**33390 [30901;34590]**
All Biallelic SNPs	**31770 [28811;38933]**	**37490 [33674;38933]**	**30460 [29534;32871]**	**30400 [29014;32652]**	**31040 [29585;33864]**	**31150 [28811;32262]**
All Biallelic INDELs	**265 [171;597]**	**452 [299;597]**	**230 [187;305]**	**217 [171;299]**	**247 [199;366]**	**239 [191;309]**
All Multiallelics	**2036 [1821;2565]**	**2383 [2140;2565]**	**1948 [1844;2255]**	**1958 [1837;2265]**	**1991 [1849;2281]**	**1999 [1821;2169]**

* Protein truncating



Supplementary Figure 8 | Single variant analyses for exome sequence and combined data sets. QQ plots for each of the three minor allele frequency categories (common, low-frequency, and rare) for (a) each of the five major ancestry groups included in the exome sequencing study (African American N=2,074; East Asian N=2,165; European N=4,541; Hispanic N=1,943; South Asian N=2,217); (b) the combined exome sequencing results ("13K Meta", N=12,940); (c) all exome array data ("Exome-array", N=79,854); and (d) exome array data combined with exome sequence data ("Combined",N=92,794). The grey region on each plot represents the (analytically estimated) 95% confidence interval.

Supplementary Table 9A | Distribution of mean age-of-diagnosis by *PAX4* Arg192His (rs2233580) genotypes in replication studies.

Study	No. of cases	Age of diagnosis (years) [Mean ± SD]			$\hat{\beta}$(SE)	*p*-value[a]
		CC	CT	TT		
SNUH	570	50.81 ± 9.67	51.61 ± 11.52	51.20 ± 10.90	0.66	0.458
Hong-Kong	489	36.41 ± 9.49	36.80 ± 10.13	32.14 ± 6.58	0.70	0.351
Singapore	560	57.12 ± 12.55	56.47 ± 13.03	52.40 ± 12.18	-1.26	0.232
Combined	1,619				0.24	0.640

$\hat{\beta}$: Regression coefficient estimates. SE: standard error
[a] Linear regression *p-value* of age-of- diagnosis with Arg192His

Supplementary Table 9B | Single-variant T2D association analysis of *PAX4* Arg192His (rs2233580) in each ancestry group from exome-sequence analysis and replication.

	Genotype counts (CC/TC/TT)		Odds ratio (95% CI)	*p*-value
Ancestral group	**Cases**	**Controls**		
Within ancestry analysis				
African American	1018 / 0 / 0	1056 / 0 / 0	-	-
East Asian	779 / 201 / 32	981 / 167 / 5	1.79 [1.47-2.19]	9.26×10^{-9}
European	2359 / 0 / 0	2182 / 0 / 0	-	-
Hispanic	1021 / 0 / 0	921 / 1 / 0	-	-
South Asian	1093 / 1 / 0	1122 / 1 / 0	-	-
Meta-analysis	-	-	1.79 [1.47-2.19]	9.26×10^{-9}
Replication studies				
SNUH	500 / 107 / 15	390 / 43 / 8	1.62 (1.20 – 2.19)	0.00186
Hong-Kong	315 / 153 / 22	260 / 80 / 3	1.76 (1.33 – 2.32)	6.81×10^{-5}
Singapore	504 / 157 / 16	568 / 147 / 10	1.24 (0.99 – 1.54)	0.0588
Meta-analysis	-	-	1.47 (1.26 – 1.70)	5.87×10^{-7}

Supplementary Table 9C | Study information

Ancestry	Study	Citation(s)	PubMed ID(s)	T2D Case Ascertainment	T2D Control Ascertainment	T1D and MODY exclusion criteria	Genotyping and QC
East Asian	*Hong Kong Diabetes Registry + Using resequencing and bioinformatics to discover a genomic signatory to predict young onset and familial type 2 diabetes [Hong Kong Chinese]*	*Ma RC, et al. Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. Diabetologia 2013; 56(6): 1291-305.*	23532257	*- Clinically diagnosed with T2D - Cases with early onset diabetes were preferentially selected*	*- Fasting blood glucose < 6.1 mmol/l - No history of diabetes - BMI ≤ 25 - No central obesity*	*- Cases with T1D presentation were excluded - Cases need insulin within 1 year of diagnosis were excluded*	*Sanger Sequencing Call Rate = 0.99 HWE P = 0.806*
East Asian	Seoul National University Hospital Diabetes Case Control Study [Korean]	*Cho, Y. S. et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet. 11;44(1):67-72 (2011)*	*22158537*	*- Clinically diagnosed as T2D using World Health Organization criteria - Participants with at least one sibling with T2D were preferentially included*	- Age ≥ 60 years old - No previous history of diabetes - No family history of diabetes in first degree relatives - Fasting plasma glucose < 6.1 mmol/l - HbA1c < 5.8%	- Diabetes patients positive for GAD antibodies were excluded - Diabetes patients with onset before age of 25 or suspected MODY cases were excluded	TaqMan genotyping Call rate =0.997 HWE P =2.0e-4 (re-sequencing confirms genotype calls of minor allele homozygotes)
East Asian	*Singapore Diabetes Cohort Study and Singapore Prospective Study Program [Singapore Chinese]*	*Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)*	*21490949*	*- Clinically ascertained T2D from primary care clinics*	*- Fasting blood glucose <6 mmol/l - No personal history of diabetes - No anti-diabetic medication*	*- Clinical records were extracted from primary care clinics, and suspected T1D and MODY cases were excluded*	*Illumina Human1M-Duo v3.0 Call rate = 1.00 HWE P = 0.86*

Supplementary Table 9D | Sample characteristics

			Cases					Controls			
Ancestry	Study	N Total	N Case	# Females (%)	Mean age (SD), years	Mean BMI (SD), kg/m^2	Mean age of diagnosis (SD), years	N Control	# Females (%)	Mean age (SD), years	Mean BMI (SD), kg/m^2
East Asian	*Hong Kong Diabetes Registry + Using resequencing and bioinformatics to discover a genomic signatory to predict young onset and familial type 2 diabetes [Hong Kong Chinese]*	*833*	*490*	*274 (55.9)*	*F: 43.5 (11.4)* *M: 46.6 (12.3)*	*F: 25.0 (4.2)* *M: 24.8 (4.1)*	*F: 34.7 (8.8)* *M: 38.4 (10.2)*	*343*	*183 (52.7)*	*F: 43.3 (9.0)* *M: 43.9 (10.5)*	*F: 22.9 (3.9)* *M: 23.9 (3.1)*
East Asian	*Seoul National University Hospital Diabetes Case Control Study [Korean]*	*1063*	*622*	*344 (55.3)*	*F:60.6 (9.6)* *M: 59.9 (10.0)*	*F: 24.6 (4.4)* *M: 24.1 (3.3)*	*F: 51.4 (9.5)* *M: 50.4 (10.7)*	*441*	*248 (56.2)*	*F: 64.4 (3.1)* *M: 65.2 (3.8)*	*F: 24.0 (3.3)* *M: 22.5 (3.6)*
East Asian	*Singapore Diabetes Cohort Study and Singapore Prospective Study Program [Singapore Chinese]*	1402	677	232(16.5)	F: 65.6(10.3) M: 66.3(9.9)	F: 25.4(4.2) M: 25.2(3.6)	F: 57.1(12.9) M:56.7(12.6)	725	281(20.0)	42.1(7.1) 44.2(8.8)	21.9(3.5) 23.4(3.2)

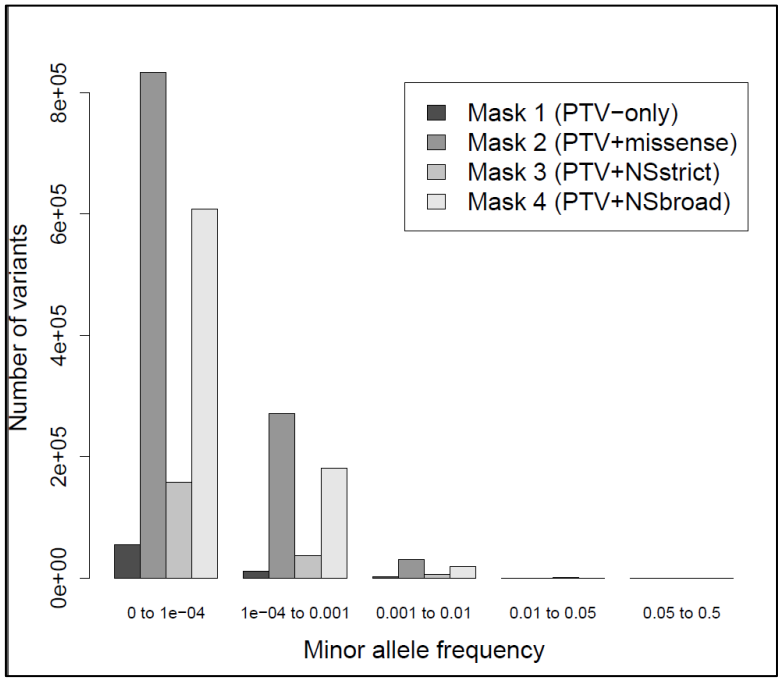
Supplementary Table 10 | Summary of biological knowledge for genes described in the paper

Gene	Biological function	Previously disease association	Mouse knockout model
ABCC8	"ATP-binding cassette sub-family C member 8". Member of the superfamily of ATP-binding cassette (ABC) transporters, which modulates ATP-sensitive potassium channels and insulin release from beta cells. Receptor for sulfonylurea antidiabetic agents (SUR1).	T2D (OMIM 125853). Leucine-sensitive hypoglycemia of infancy (OMIM 240800). Familial hyperinsulinemic hypoglycemia 1 (OMIM 256450). Transient neonatal diabetes mellitus type 2 (OMIM 610374). Permanent neonatal diabetes mellitus (OMIM 606176).	Adult mice are glucose intolerant with reduced glucose-stimulated insulin secretion from isolated islets (PMID 10734066).
COBLL1	"Cordon-bleu protein-like 1". Contains two WH2 (Wiskott-Aldrich homology 2) actin binding domains.	T2D (PMID 23160641). Metabolic Syndrome (PMID 24981077). Lower fasting insulin and lower insulin resistance in obese children (PMID 23463496). BMI-adjusted waist-hip-ratio in women (PMID 23754948).	No knockout reported.
FES	"Tyrosine-protein kinase Fes/Fps". Cytoplasmic protein tyrosine kinase implicated in the regulation of the actin cytoskeleton and myeloid differentiation. Promotes survival of leukemia cells and is present as an activated kinase in some patients with acute myeloid leukemia.	Feline sarcoma (Fes) / Fujinami avian sarcoma (Fps) viral oncogene homolog (OMIM 190030).	No metabolic phenotypes identified. Knockouts show partial lethality and runting. Adults have abnormal hematopoiesis and leukocyte function (PMID 11021537).
GCKR	"Glucokinase (hexokinase 4) regulator". Member of Sugar Isomerase (SIS) family of proteins that forms a complex with and inhibits glucokinase in liver and possibly pancreatic beta cells and neurons.	T2D (OMIM 125853). Serum glucose and insulin-related traits (PMID 20081858). Serum triglycerides (PMID 18193043) and LDL cholesterol (PMID 18179892). Plasma C-reactive protein (PMID 18439548).	Reduced hepatic glucokinase expression and activity. Altered glucose and insulin homeostasis worsened by high-sucrose diet (PMID 10588736).
GPSM1	"G-protein-signaling modulator 1". Receptor-independent activator of basal G-protein signaling. Inhibits GDP-dissociation from Galpha(i). Associated SNP affects a putative interaction site for STK11/LKB1	T2D (PMID 23945395).	Reduced body weight and white adipose tissue mass. Increased food consumption and increased nocturnal energy expenditure. Altered blood pressure homeostasis (PMID 18450958).
GRB14	"Growth factor receptor-bound protein 14" Adapter protein containing SH2, ras-associating and pleckstrin homology domains, involved in receptor kinase singalling. Negative regulator of insulin receptor activation of ERK1/2 (PMID 11726652). Promotes PDPK1 recruitment to activated insulin receptor and subsequent PKB/AKT activation (PMID 15210700).	T2D (PMID 21874001). Metabolic Syndrome (PMID 24981077)	Decreased body weight. Improved glucose tolerance. Reduced circulating insulin levels. Increased signaling through IRS-1 and PKB in liver and skeletal muscle (PMID 14749734).
HNF1A	"Hepatocyte nuclear factor 1-alpha". Transcription factor regulating expression of tissue-specific genes in hepatocytes and islet and exocrine cells of the pancreas.	T2D (OMIM 125853, OMIM 612520). Maturity-onset diabetes of the young type 3 (OMIM 600496).	Reduced serum insulin and increased serum glucose levels. Decreased linear growth with GH resistance. Hepatic steatosis with abnormal liver function (PMID 9566924, PMID 8598044).
HNF4A	"Hepatocyte nuclear factor 4-alpha". Nuclear receptor involved in regulation of liver-specific transcripts, including genes involved in gluconeogenesis and fat metabolism (PMID 23485969).	T2D (OMIM 125853). Maturity-onset diabetes of the young type 1 (OMIM 125850).	knockout shows premature death (by early adulthood) with abnormal liver morphology and function. Isoform-specific knockouts show impaired glucose tolerance (HNF4alpha1) and dyslipidemia (HNF4alpha7) (PMID 11158324, PMID 16498401).
KCNJ11	"ATP-sensitive inward rectifier potassium channel 11". Potassium inwardly-rectifying channel. Activity regulated by G-proteins. Provides channel pore whose activity is regulated by the sulfonylurea receptor ABCC8/SUR1 in the beta-cell ATP-sensitive potassium channel.	T2D (OMIM 125853). Familial hyperinsulinemic hypoglycemia type 2 (OMIM 601820). Transient neonatal diabetes mellitus type 3 (OMIM 610582). Permanent neonatal diabetes mellitus (OMIM 606176).	Metabolic phenotypes include impaired insulin secretion and mild glucose intolerance. Reduced glucagon secretion in response to hypoglycemia. Increased energy expenditure with reduced susceptibility to diet-induced obesity (PMID 11319559, PMID 20074528).
MTMR3	"Myotubularin-related protein 3". Myotubularin dual specificity protein phosphatase. Phosphatase can hydrolyze phosphatidylinositol 3-phosphate, phosphatidylinositol 3,5-bisphosphate and other phosphoinositide lipids.	LDL cholesterol (PMID 24097068). Lung cancer (PMID 21725308). Early-onset inflammatory bowel disease (PMID 19915574): IgA nephropathy (PMID 21725308).	Impaired glucose tolerance (males). Increased serum alkaline phosphatase (both sexes only) (JAX J103485).
PAM	"Peptidyl-glycine alpha-amidating monooxygenase". Catalyzes C-terminal alpha-amidation of peptides, which is required for full bioactivity of some neuropeptides and peptide hormones. Variant alters intragranular domain in NHL2 repeat of peptidyl-alpha-hydroxyglycine alpha-amidating lyase region.	T2D (PMID 24464100). Insulinogenic index (PMID 23263489).	Knockout mice die in mid-gestation (e14.5-e15.5). Old heterozygotes (age >10 months) have mild glucose intolerance and increased white fat mass (PMID 16225857).
PAX4	"Paired box protein Pax-4". Transcriptional repressor containing homeobox domain that binds to a common element in the glucagon, insulin and somatostatin promoters. Critical roles during fetal pancreatic islet development and differentiation of insulin producing beta cells. Variant alters highly conserved amino acid in homeodomain region (Arg192His) and is predicted to be deleterious by some scores. Located in the GCC1 locus identified by T2D GWA Studies.	T2D (OMIM 125853). Maturity-onset diabetes of the young type 9 (OMIM 612225). Ketosis-prone Diabetes Mellitus (PMID 612227).	Knockout mice show early postnatal lethality (within 3d of birth). Pancreatic islets lack cells expressing insulin and somatostatin and contain cells expressing ghrelin, glucagon and islet amyloid polypeptide (PMID 18058910).
PPARG	"Peroxisome proliferator-activated receptor gamma". Nuclear receptor activated by peroxisome proliferators including hypolipidemic drugs and fatty acids. Key transcriptional regulator of adipocyte differentiation. Controls expression of the peroxisomal beta-oxidation pathway of fatty acids.	T2D (OMIM 125853). Obesity (OMIM 601665). Carotid intimal medial thickness 1 (OMIM 609338). Familial partial lipodystrophy type 3 (OMIM 604367).	Adipose-specific PPARG knockout shows diminished levels of leptin and adiponectin. PPARG liver knockout shows lower cholesterol, FFA, TG. (PMID 10549291, PMID 15070754, PMID 10675354).
PPIP5K2	"Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase 2". Bifunctional inositol kinase that acts with the IP6K kinases to synthesize diphosphate group-containing inositol pyrophosphates including diphosphoinositol pentakisphosphate. Regulates a variety of cellular processes including vesicle trafficking, exocytosis, insulin signaling and apoptosis.	No disease associations.	No knockout reported.
RREB1	"Ras-responsive element-binding protein 1". Krueppel C2H2-type zinc-finger protein that binds specifically to the RAS-responsive elements of gene promoters. Variant alters a conserved residue (Asp1171Asn) predicted to be deleterious by some algorithms.	Body fat distribution (PMID 23966867). Serum urate levels (PMID 20884846).	No knockout reported.
SLC30A8	"Solute carrier family 30 member 8". Zinc efflux transporter involved in the accumulation of zinc in intracellular vesicles. Colocalizes with insulin in the secretory pathway granules of insulin-secreting cells. Highly expressed in pancreatic islets.	T2D (ClinVar=RCV000001055.1).	Reduced islet zinc content. abnormal beta-cell morphology on EM. Inconsistent effects on insulin processing and glucose-induced insulin secretion in vitro. Increased hepatic insulin clearance. Intraperitoneal glucose tolerance tests show impaired glucose tolerance in young mice (4-6 weeks) (PMID 24051378, PMID 24751356).
THADA	"Thyroid adenoma-associated protein". Locus is disrupted by chromosomal rearragnements involving 2p21 in some benign thyroid adenomas (PMID 12955091). Associated with lower beta-cell response to GLP-1 and arginine (PMID 19833888). Ubiquitously expressed transcript.	T2D (OMIM 125853). Crohn's disease (PMID 21102463). Prostate cancer (PMID 19767753). Multiple Sclerosis (PMID 22190364). Polycystic ovary syndrome (PMID 21151128).	No knockout reported.
TM6SF2	"Transmembrane 6 superfamily member 2". Gene located in the CILP2-TMS6F2 locus identified by T2D GWA Studies. Ubiquitously expressed transcript encodes endoplasmic reticulum protein that regulates fat metabolism in liver (PMID 24927523).	T2D (PMID 22885922). Glu167Lys variant associated with increased susceptibility to nonalcoholic fatty liver disease (PMID 24531328), increased total cholesterol and myocardial infarction risk (PMID 24633158).	No knockout reported. Adeno-associated virus-mediated shRNA knockdown increased hepatic triglyceride content and decreased very-low-density lipoprotein secretion (PMID 24531328).
TSPAN8	"Tetraspanin-8". Member of the transmembrane 4 glycoprotein superfamily, which is known to interact with integrins.	SNPs (PMID 18372903) and CNV (PMID 20360734) near TSPAN8 associated with T2D.	Lower total body weight, lower fat and lean mass (males only). Lower bone mineral density and phosphorus levels (males only) (PMID 20733586).
WFS1	"Wolframin". Endoplasmic reticulum protein involved in cellular Ca++ homeostasis. Functions in pancreatic beta cell and neuronal survival. Mutations can cause Wolfram Syndrome (also known as DIDMOAD - Diabetes Insipidus. Diabetes Mellitus. Optic Atrophy and Deafness).	T2D (PMID 17603484). Wolfram syndrome 1 (OMIM 222300). Autosomal dominant deafness 6 (OMIM 600965)	Reduced beta cell mass. Activated ER-stress response associated with apoptosis in beta-cells. Increased blood glucose and decreased insulin secretion (PMID 15056606, PMID 16215705, PMID 19041897).

Supplementary Table 11a | Gene-level mask descriptions. Protein truncating (PTV) and missense variants were further annotated to identify variants predicted deleterious by at least one (NS_{broad}) or each of five (NS_{strict}) prediction algorithms (LRT, Mutation Taster, PolyPhen2-HumDiv, PolyPhen2-HumVar, SIFT). PTV, missense, NS_{strict}, and NS_{broad} classes of variants were combined to generate four masks for gene-level testing (described in first column). The second column lists variant categories (and variant counts) contributing to each mask; the third column indicates the total numbers of variants in each mask; the fourth column indicates the number of genes containing at least one variant meeting mask criteria; the final column indicates the median number (and range) of variants per gene for each mask.

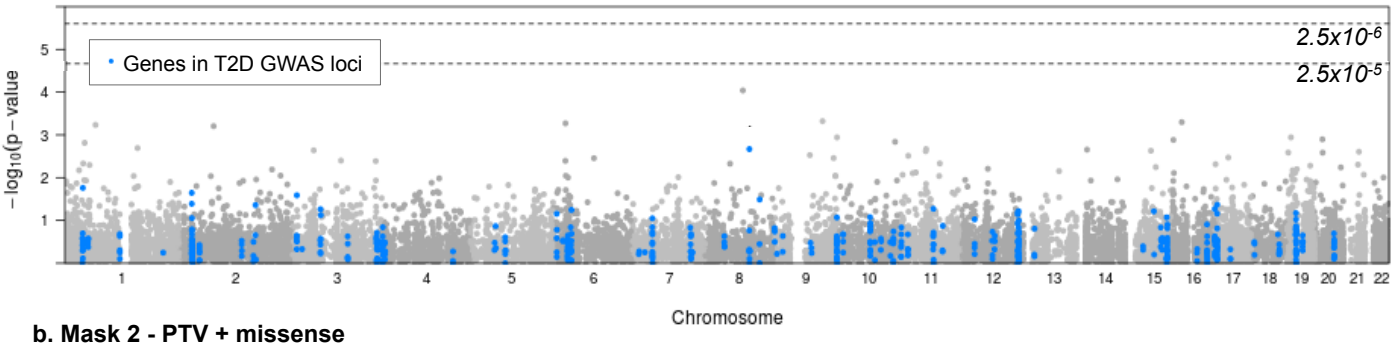
Mask name and description	Variant classes (number of variants)	Number of variants	Number of genes	Median number of variants/gene [range]
1) Mask 1 - PTV-only *Predicted protein-truncating variants (PTVs)*	PTVs (69,956)	69,956	14,415	3 [1-135]
2) Mask 2 - PTV + missense *PTVs and missense variants with MAF<1%, as high-frequency variants may be less likely to be deleterious*	PTVs (69,956) missense, MAF<1% (1,065,607)	1,135,563	17,806	47 [1-4284]
3) Mask 3 - PTV + NS_{strict} *PTVs and missense variants predicted deleterious by all five algorithms (NS_{strict}): LRT, Mutation Taster, PolyPhen2-HumDiv, PolyPhen2-HumVar, SIFT*	PTVs (69,956) NS_{strict} (131,976)	201,932	16,757	8 [1-429]
4) Mask 4 - PTV + NS_{broad} *PTVs and NS_{strict} variants, plus missense variants predicted deleterious by at least one algorithm (NS_{broad}) and with MAF<1%*	PTVs (69,956) NS_{strict} (131,976) NS_{broad}, MAF<1% (603,369)	805,301	17,771	33 [1-2124]

Supplementary Table 11b | Numbers of variants for each mask in 12,940 WES samples.

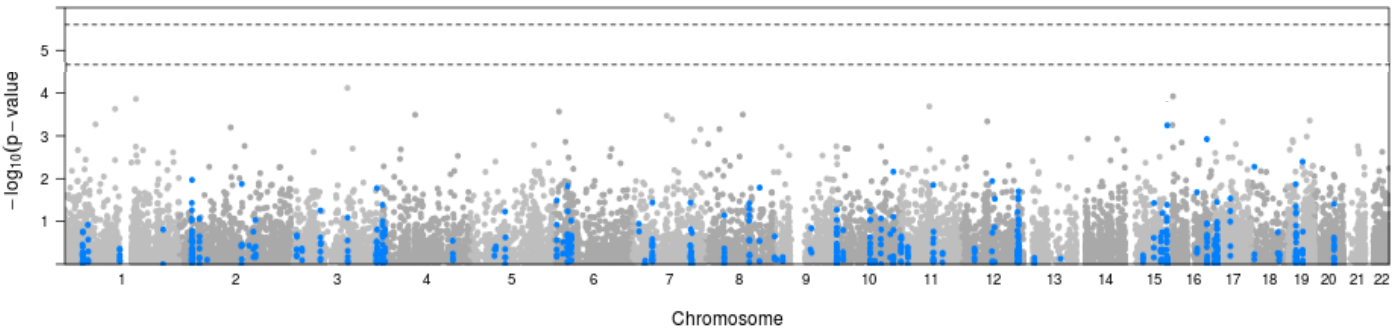


Supplementary Figure 12 | Manhattan plots for gene-level analysis in 12,940 WES samples

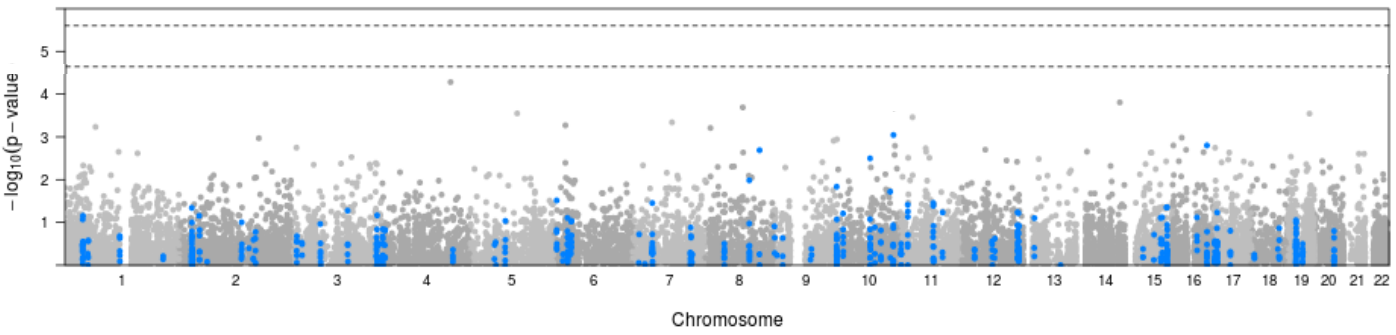
a. Mask 1 - PTV-only



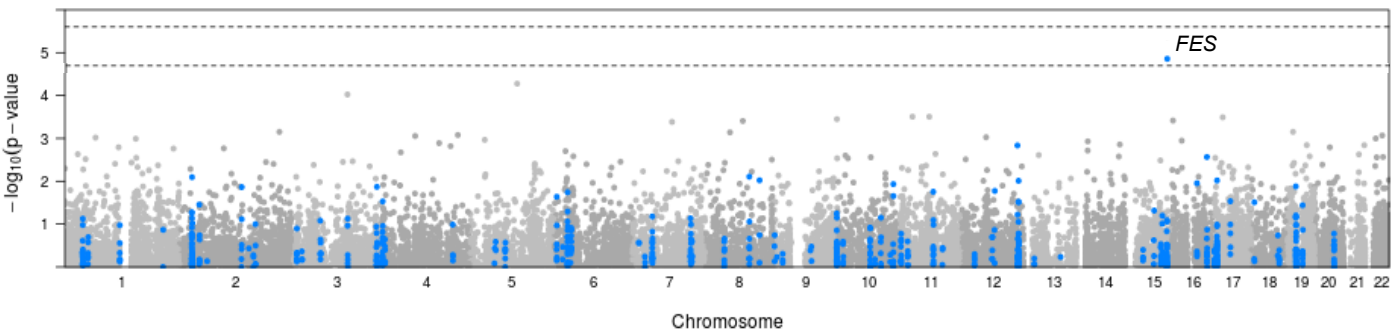
b. Mask 2 - PTV + missense

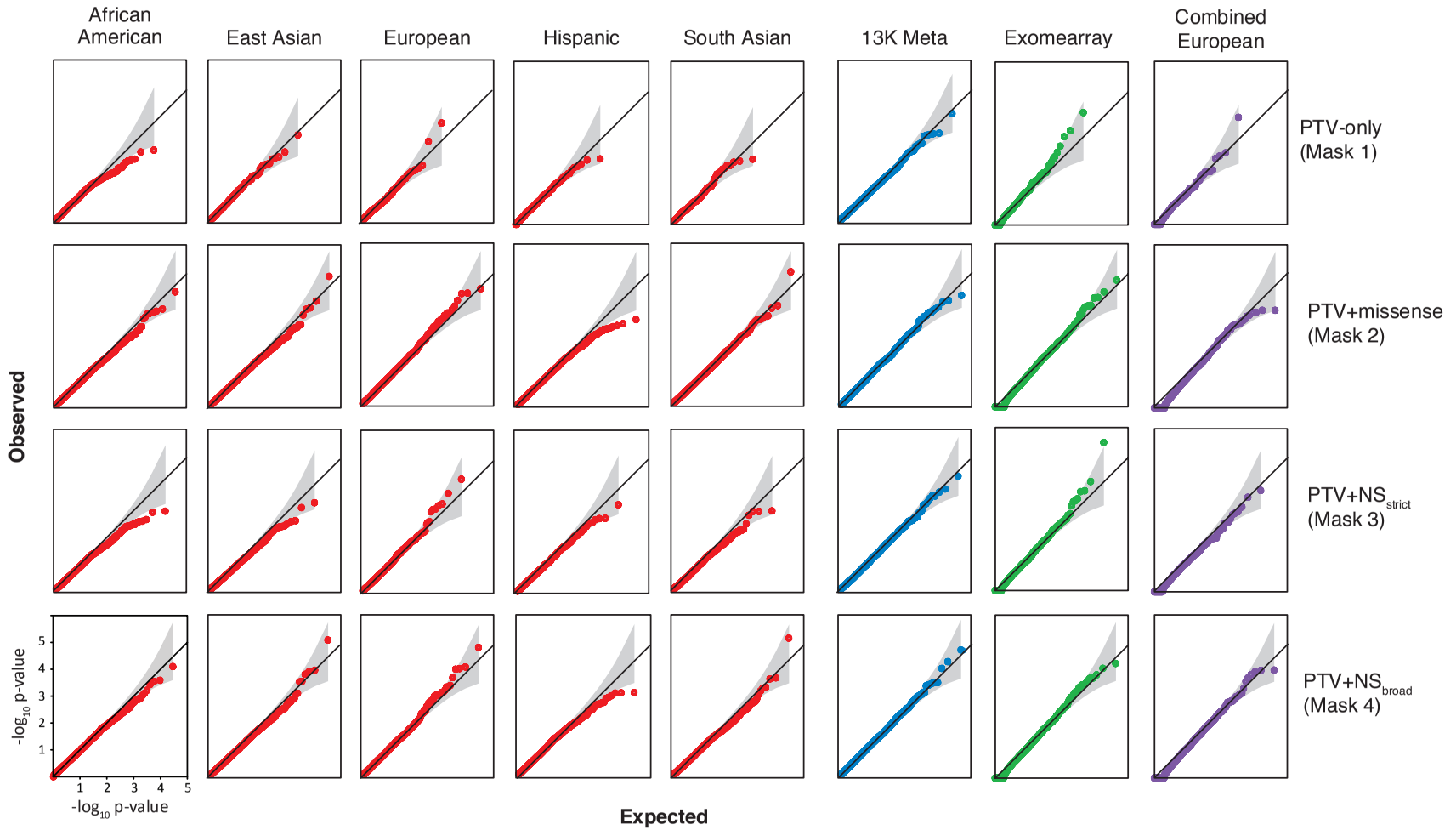


c. Mask 3 - PTV + NS_{strict}



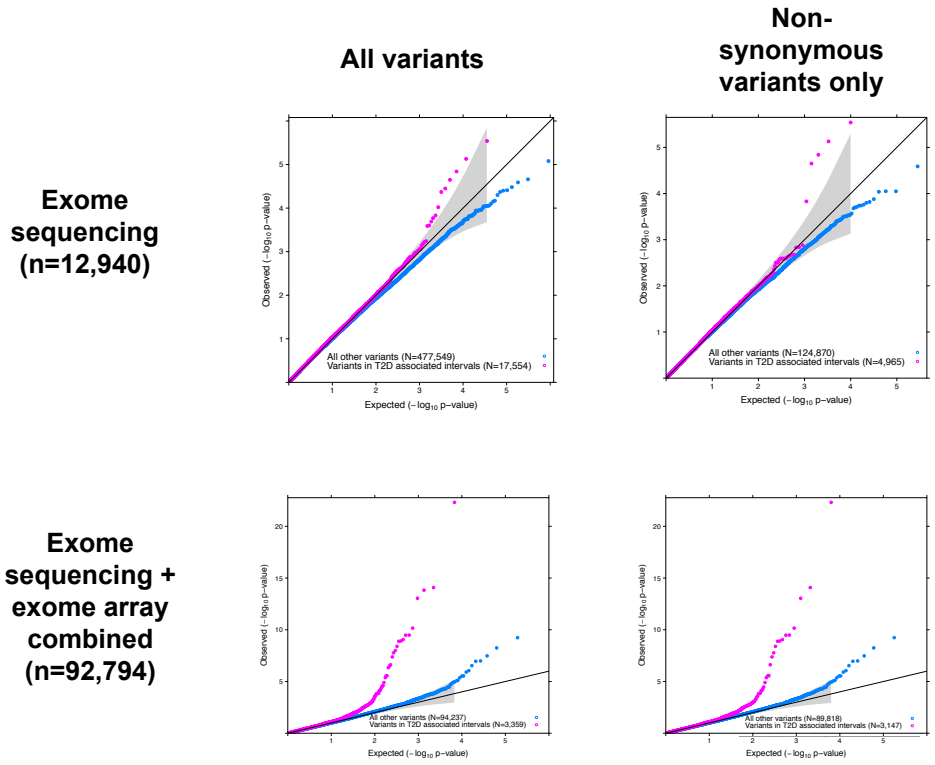
d. Mask 4 - PTV + NS_{broad}





Supplementary Figure 13 | Aggregate (gene-based) analyses for exome sequence and combined data sets. QQ plots for each of the four variant masks for (a) each of the five major ancestry groups in the exome sequencing study (African American N=2,074; East Asian N=2,165; European N=4,541; Hispanic N=1,943; South Asian N=2,217); (b) the combined exome sequencing results ("13K Meta", N=12,940); (c) all exome chip data ("Exome array", N=79,854); and (d) exome chip data combined with exome sequence data from Europeans only ("Combined European", N=84,395). The grey region on each plot represents the (analytically estimated) 95% confidence interval. Across all analyses, there is no compelling evidence that results depart from the null.

Supplementary Figure 14A | Single variant analyses in GWAS regions. The QQ plots display single variant analyses for all variants (left) and nonsynonymous variants only (right). Analyses of exome sequence (6,504 cases; 6,436 controls) are in the upper panels, and of the combination of exome sequence and exome array data (34,809 cases, 57,985 controls) in the lower. In each panel, variants mapping to established GWAS regions are in pink, and all other variants in blue (only variants with a minor allele count over 9 are included). The plots show enrichment of association signals for coding variants in established GWAS signals resulting from a combination of linkage disequilibrium to known common variant GWAS signals, and secondary signals at a subset of loci (eg *HNF4A, THADA, TSPAN8*).



Supplementary Table 14B | *FES* gene-level association statistics for all ancestry groups (African American N=2,074; East Asian N=2,165; European N=4,541; Hispanic N=1,943; South Asian N=2,217; Total N=12,940) for Mask 4 (PTV + NS_{broad}).

Samples	No. variants; MAC control / case	Max. single-marker MAC (%)	SKAT-O p-value	Top single-marker p-value
Afr. Amer.	23; 25/25	12 (24)	0.83	0.13
E. Asian	18; 15.2/10	3 (11.9)	0.68	0.11
European	24; 21/17	6 (15.8)	0.51	0.062
Hispanic	12; 5/10.5	3 (19.4)	0.22	0.068
S. Asian	21; 19/45	38 (59.4)	7.2×10^{-6}	7.5×10^{-6}
All	81; 85.3/107.6	38 (19.7)	1.9×10^{-5}	7.5×10^{-6}

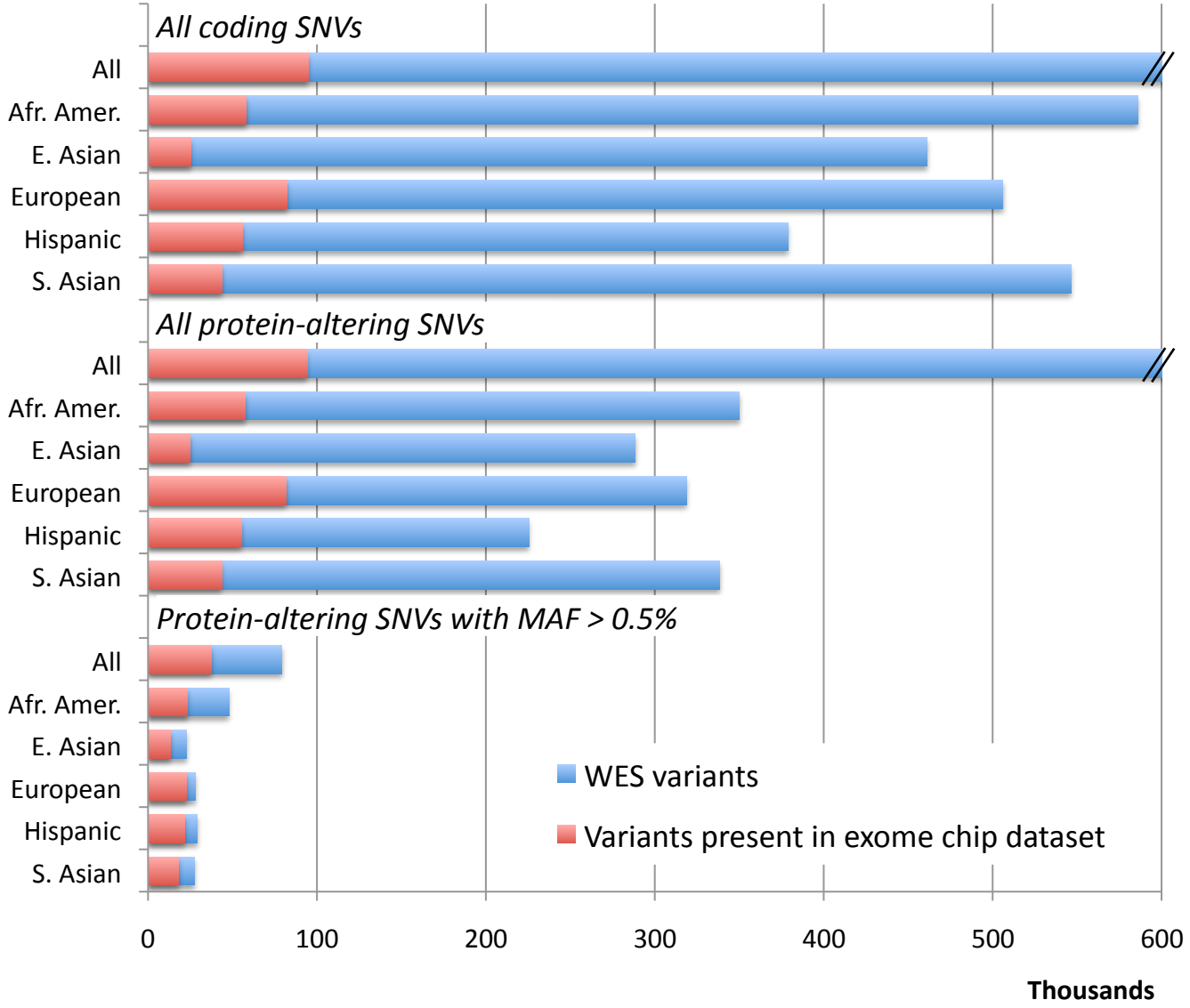
Supplementary Table 15a | Exome array cohort information

Ancestry	Study	Citation(s)	PubMed ID(s)	T2D Case Ascertainment	T2D Control Ascertainment	T1D and MODY exclusion criteria	Genotyping array	Calling algorithm	Association covariates
European	Oxford-based UK T2D case-control	Voight BF et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010 Jul;42(7):579-89. Spector TD, Williams FM.The UK Adult Twin Registry (TwinsUK). Twin Reserch and Human Genetics 2006; 9 (6): 899-906. Tan GD, Neville MJ, Liverani E, Humphreys SM, Currie JM, Dennis L, Fielding BA, Karpe F. The in vivo effects of the Pro12Ala PPARgamma2 polymorphism on adipose tissue NEFA metabolism: the first use of the Oxford Biobank. Diabetologia. 2006 Jan;49(1):158-68. Strachan DP, et al. Lifecourse influences on health among British adults: Effects of region of residence in childhood and adulthood. Int. J. Epidemiol. 2007;36:522-531.	20581827, 17254428, 16362285, 17255346	The T2D cases were selected from UK Caucasian subjects who are part of the Diabetes UK Warren 2 repository. The remainder were recruited as isolated cases but these cases were (compared to population-based cases) of relatively early onset and had a high proportion of T2D parents and/or siblings. T2D was defined as current prescribed treatment with sulphonylureas, biguanides, other oral agents and/or insulin or, in the case of individuals treated with diet alone, historical or contemporary laboratory evidence of hyperglycemia.	Controls from all sources were selected without reference to T2D status and fasting glucose ≤7.0 mmol/l.	Individuals with maturity-onset diabetes of the young and mitochondrial diabetes, were excluded. Other inclusion criteria included: absence of first-degree relatives with type 1 diabetes; an interval of ≥1 year between diagnosis and institution of regular insulin therapy; and negative testing for antibodies to glutamic acid decarboxylase (anti-GAD). An anti-GAD titer >10 U (corresponding to ~8 SD above the mean of 88 normal control subjects) in duplicate samples was considered positive.	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + zCall	
European	The Diabetes Audit and Research in Tayside Scotland (GoDarts)	Morris AP et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012 Sep;44(9):981-90.	22885922	Cases had T2D diagnosed between the ages of 35-70 years (inclusive). The diagnosis of diabetes was based on either current prescribed treatment with diabetes-specific medication or, in the case of individuals treated with diet alone, laboratory evidence of diabetes as defined by the WHOS8,S38.	Controls were defined as having no diagnosis of diabetes at the time of recruitment (or subsequently), fasting glucose ≤7.0 mmol/l, HbA1c ≤6.4% and age < 80 years.	Cases were excluded if they had an established (clinical and/or molecular) diagnosis of monogenic diabetes (e.g. maturity-onset diabetes of the young, mitochondrial diabetes) or if they had been treated with regular insulin therapy within 1 year of diagnosis. No islet autoantibodies were measured.	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + zCall	
European	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	Lind L et al. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. Arterioscler Thromb Vasc Biol 2005 25 2368-2375	16141402	Known T2D or fasting whole blood glucose ≥ 6.1mmol/l	Controls were defined as having no diagnosis of diabetes at the time of recruitment (or subsequently), whole blood glucose ≤6.0 mmol/l.	-	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + zCall	
European	Uppsala Longitudinal Study of Adult Men (ULSAM)	Ingelsson, E. et al. Insulin resistance and risk of congestive heart failure. Jama 294, 334-41 (2005).	16030278	Hospital discharge register-defined diabetes before 2002	Controls were defined as having no diagnosis of diabetes at the time of recruitment.	-	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + zCall	
European	Metabolic Syndrome in Men Study (METSIM)	Stancakova, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes 58, 1212–1221 (2009)	19223598	T2D as classified by WHO 1997 criteria (fasting plasma glucose >= 7.0 mmol/l or 2-hr plasma glucose >= 11.1 mmol/l)	NGT as classified by WHO 1997 criteria (fasting plasma glucose < 6.1 mmol/l and 2-hr plasma glucose < 7.8 mmol/l)	known T1D cases excluded	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + manual review	age, batch
European	FIN-D2D 2007	Kotronen A, Yki-Järvinen H, Männistö S, Saarikoski L, Korpi-Hyövälti E, Oksa H, Saltevo J, Saaristo T, Sundvall J, Tuomilehto J, Peltonen M. Non-alcoholic and alcoholic Fatty Liver Disease – two Diseases of Affluence associated with the Metabolic Syndrome and Type 2 Diabetes: the FIN-D2D Survey. BMC Public Health 2010; 10: 237.	20459722	T2D as classified by WHO 1999 criteria (fasting plasma glucose >= 7.0 mmol/l or 2-hr plasma glucose >= 11.1 mmol/l)	NGT as classified by WHO 1999 criteria (fasting plasma glucose < 6.1 mmol/l and 2-hr plasma glucose < 7.8 mmol/l)	known T1D cases excluded	Illumina HumanExome-12v1-1_A Beadchip	Illumina GenCall using standard Illumina cluster files + Zcall	age, sex
European	The Dose Responses to Exercise Training (DR's EXTRA) Study	Diet, fitness and metabolic syndrome - The DR's EXTRA Study. Kouki R, Schwab U, Lakka TA, Hassinen M, Savonen K, Komulainen P, Krachler B, Rauramaa R. Nutr Metab Cardiovasc Dis. 2012 Jul;22(7):553-60. Epub 2010 Dec 24	21186108	T2D as classified by WHO 1999 criteria (fasting plasma glucose >= 7.0 mmol/l or 2-hr plasma glucose >= 11.1 mmol/l) or physician diagnosed	NGT as classified by WHO 1999 criteria (fasting plasma glucose < 6.1 mmol/l and 2-hr plasma glucose < 7.8 mmol/l)	known T1D cases excluded	Illumina HumanExome-12v1-1_A Beadchip	Illumina GenCall using standard Illumina cluster files + Zcall	age, sex
European	National FINRISK 2007 Study (FINRISK 2007)	Thirty-five-year trends in cardiovascular risk factors in Finland. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Männistö S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P. Int J Epidemiol. 2010 Apr;39(2):504-18	19959603	T2D as classified by WHO 1999 criteria (fasting plasma glucose >= 7.0 mmol/l or 2-hr plasma glucose >= 11.1 mmol/l)	NGT as classified by WHO 1999 criteria; frequency matched to cases by birth province; BMI ≥18.5 kg/m2; within each birth province, prioritized samples with highest values for age + 2*BMI	known T1D cases excluded	Illumina HumanExome-12v1-1_A Beadchip	Illumina GenCall using standard Illumina cluster files + Zcall	age, sex
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Valle, T. et al. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. Diabetes Care 21(6), 949-958 (1998); Scott, L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316(5829), 1341-1345 (2007)	9614613; 17463248	T2D as classified by WHO 1999 criteria, by report of diabetes medication use, or based on medical record review	NGT as classified by WHO 1999 criteria; approximately frequency matched to the cases by 5-year age category, sex, and birth province.	known T1D cases excluded	Illumina HumanExome-12v1-1_A Beadchip	Illumina GenCall using standard Illumina cluster files + Zcall	age, sex
European	Prevalence, Prediction and Prevention of diabetes (PPP) (Finnish)	Isomaa B, Forsén B, Lahti K, Holmström N, Wadén J, Matintupa O, Almgren P, Eriksson JG, Lyssenko V, Taskinen MR, Tuomi T, Groop LC. A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study. Diabetologia. 2010 Aug;53(8):1709-13.	20454776	Diagnosis of diabetes was based on an OGTT or a history of previously known diabetes applying WHO criteria. In uncertain cases, the diagnosis was confirmed from patient records.	No previously known diabetes. Free from diabetes after an OGTT applying WHO criteria.		Illumina HumanExome-12v1_A Beadchip	Custom birdseed algorithm within batch	age, sex, analysis performed within batch
European	Diabetes Registry Vaasa (DIREVA) (Finnish)			- Previous diagnosis of T2D - Normal C-peptide levels - No Anti-GAD antibody	No controls are in the registry	No Anti-GAD antibody and normal C-peptide levels	Illumina HumanExome-12v1_A Beadchip	Custom birdseed algorithm within batch	age, sex, analysis performed within batch
European	All New Diabetics In Scania (ANDiS) (Swedish)			- Previous diagnosis of T2D - Normal C-peptide levels - No Anti-GAD antibody	No controls are in the registry	No Anti-GAD antibody and normal C-peptide levels	Illumina HumanExome-12v1_A Beadchip	Custom birdseed algorithm within batch	age, sex
European	Malmö Diet and Cancer (MDC) (Swedish)	Manjer J, Carlsson S, Elmståhl S, Gullberg B, Janzon L, Lindström M, Mattisson I, Berglund G. The Malmö Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev 2001;10:489-99. and Manjer J, Elmståhl S, Janzon L, Berglund G. Invitation to a population-based cohort study: differences between subjects recruited using various strategies. Scand J Public Health 2002;30:103-12.	11916347 12028859	DM at baseline was defined as self-report of a physician diagnosis or use of diabetes medication or fasting whole blood glucose greater than or equal to 6.1 mmol/l (corresponding to fasting plasma glucose concentration >= 7.0 mmol/l).	Fasting blood glucose below 6.1 mmol/l (corresponding to fasting plasma glucose concentration < 7.0 mmol/l).	None	Illumina HumanOmniExpressExome-8v1_B	Custom birdseed algorithm within batch	age, sex, analysis performed within batch
European	Scania Diabetes Registry (SDR) (Swedish)	Lindholm E, Agardh E, Tuomi T, Groop L, Agardh C.D. Classifying diabetes according to the new WHO clinical stages. Eur J Epidemiol 2001;17, 983-9.	12380709	Physicians own classification into T2D, based on WHO 1985 guidelines (before 2001) or WHO 1999 guidelines (diagnosed after January 2001).	No controls are in the registry	Presenve of severe hyperglycaemia and/or ketosis at diagnosis, low fasting C-peptide levels and presence or GAD antibodies	Illumina HumanExome-12v1_A Beadchip	Custom birdseed algorithm within batch	age, sex, analysis performed within batch
European	Nurses' Health Study (NHS)	Qi L et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Human Molecular Genetics 2010;19(13):2706-15	20418489	Diabetes cases were defined as self-reported diabetes confirmed by a validated supplementary questionnaire. For cases before 1998, diagnosis was made using criteria consistent with those proposed by the National Diabetes Data Group. We used the American Diabetes Association diagnostic criteria for diagnosis of diabetes cases during the 1998 and 2000 cycles. A 98% of self-reported cases were confirmed by medical records review.	Controls were defined as those free of diabetes at the time of diagnosis of the case and remained unaffected through follow-up	NA	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + manual review	age
European	Health Professional Follow-Up Study (HPFS)	Qi L et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Human Molecular Genetics 2010;19(13):2706-15	20418489	Diabetes cases were defined as self-reported diabetes confirmed by a validated supplementary questionnaire. For cases before 1998, diagnosis was made using criteria consistent with those proposed by the National Diabetes Data Group. We used the American Diabetes Association diagnostic criteria for diagnosis of diabetes cases during the 1998 and 2000 cycles. A 98% of self-reported cases were confirmed by medical records review.	Controls were defined as those free of diabetes at the time of diagnosis of the case and remained unaffected through follow-up	NA	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + manual review	age
European	Estonian Genome Center, University of Tartu (EGCUT)	Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, Perola M, Ng PC, Mägi R, Milani L, Fischer K, Metspalu A. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol. 2014 Feb 11.	24518929	- Previous diagnosis of T2D	Population based controls with fasting glucose < 7 mmol/l	NA	Illumina HumanExome-12v1_A Beadchip	Illumina GenCall + zCALL	age, sex
European	EFSOCH and DARE	The Exeter Family Study of Childhood Health (EFSOCH): study protocol and methodology. Knight B, Shields BM, Hattersley AT.	16466435	Type 2 diabetic individuals were ascertained from the Exeter branch of the Diabetes Alliance for Research in England (DARE) study and selected for genotyping if they were not on insulin within the first year of diagnosis and diagnosed after the age of 35 years.	Control individuals were selected from the Exeter Family Study of Childhood health (EFSOCH). Male and female partners were ascertained at the time of pregnancy and included on the basis that they represent a very similar geographic distribution to the DARE case individuals, and were normoglycaemic on the basis of fasting glucose of HbA1C.	age at diagnosis 35 years of over. Not on insulin for first year of diagnosis.	Illumina HumanExome-12v1_A Beadchip	Illumina + zCall	age and sex
European	Cooperative Health Research in the Region of Augsburg [KORA]	Holle R, Happich M, Löwel H, Wichmann HE, MONICA/KORA Study Group. KORA - a research platform for population based health research. Gesundheitswesen. 2005 Aug;67 Suppl 1:S19-25. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen-resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67, Suppl. 1, S26-S30 (2005).	16032513, 16032514	- Previous diagnosis of T2D, or both fasting and 2-hr criteria met for new T2D diagnosis - Family history of diabetes (parents, sibs, children, grandparents, avuncular, cousins) - Unrelated individuals based IBS analyses	- No diagnosis of T2D - Normal glucose tolerance at baseline - Unrelated samples	-	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain + zCall	age, sex
European	Danish T2D case-control	Albrechtsen, A. et al. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia 56, 298-310 (2013)	23160641	Screen-detected (by fasting glucose or 2-hr glucose after OGTT) or clinical onset type 2 diabetes (WHO 1999 criteria)	Population-based sampled with fasting glucose < 6.1 mmol/L and 2-hr glucose < 7.8 (if measured) (WHO 1999 criteria)	Fasting C-peptide <150 pmol/L	Illumina HumanExome-12v1_A Beadchip	Illumina GenTrain version 1.0 + zCall	age, sex, (BMI)
European	GLACIER	Hallmans, G., et al., Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort - evaluation of risk factors and their interactions. Scand J Public Health Suppl, 2003. 61: p. 18-24.	14660243	- Participants with incident type 2 diabetes were identified from the Diabetes Register in Northern Sweden (DiabNorth) Web: http://www.diabetesregister.se/en/about-diabnorth-and-the-diabetes-register	- fasting glucose <5.6 mmol/l	-	Illumina HumanExome Beadchip 12 v1.1	Illumina GenTrain + zCall	age, sex, 1-10 principal components, (BMI)
European	EPIC-Norfolk (T2D cases) and the Fenland study (cohort)	Day, N. et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. British journal of cancer 80 (supp 1), 95-103 (1999); Rolfe, E. et al. Association between birth weight and visceral fat in adults. AJCN 92(2), 347-52 (2010)	10466767; 20519560	Clinically diagnosed incident cases of T2D from the EPIC-Norfolk study and prevalent undiagnosed T2D in the Fenland study based on fasting glucose >=7.0 mmol/L and/or 2hGlu >=11.1mmol/L.	Random sample of population-based Fenland study with FG<7.0mmol/L and 2h-glucose <11.1mmol/L.	NA	Illumina HumanExome-12v1_A Beadchip	Illumina GenCall + zCall	age, sex, (BMI)

Supplementary Table 15b | Exome array sample characteristics

			Cases					Controls			
Ancestry	**Study**	**N Total**	**N Case**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**	**Mean age of diagnosis (SD), years**	**N Control**	**# Females (%)**	**Mean age (SD), years**	**Mean BMI (SD), kg/m^2**
European	Oxford-based UK T2D case-control	12743	1861	767 (51.7)	F: 53.3 (11.4) M: 53.1 (10.3)	F: 34.1 (7.2) M: 31.2 (5.6)	F: NA M: NA	10882	5623 (41.2)	F: 48.5 (9.0) M: 50.7 (7.7)	F: 25.9 (5.0) M: 27.2 (4.0)
European	The Diabetes Audit and Research in Tayside Scotland (GoDarts)	3508	1715	682 (39.8)	F: 64.2 (9.4) M: 63.2 (9.4)	F: 33.3 (6.9) M: 31.5 (5.8)	F: NA M: NA	1793	820 (45.7)	F: 57.9 (11.4) M: 59.5 (11.0)	F: 26.6 (4.8) M: 27.3 (4.0)
European	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	961	111	46 (41.1)	F: 70.2 (0.1) M: 70.1 (0.1)	F: 29.9 (6.7) M: 26.6 (3.8)	F: NA M: NA	850	428 (50.3)	F: 70.3 (0.15) M: 70.1 (0.17)	F: 26.8 (4.6) M: 28.6 (3.8)
European	Uppsala Longitudinal Study of Adult Men (ULSAM)	1101	160	0 (0)	F: NA M: 71.0 (0.7)	F: NA M: 28.0 (4.0)	F: NA M: NA	941	0 (0)	F: NA M: 71.0 (0.63)	F: NA M: 25.9 (3.2)
European	Metabolic Syndrome in Men Study	5158	773	0 (0)	F: 0 (0) M: 60.7 (6.7)	F: 0 (0) M: 29.9 (5.2)	F: NA M: NA	4385	0 (0)	F: 0 (0) M: 57.1 (7.2)	F: 0 (0) M: 26.2 (3.5)
European	FIN-D2D 2007	2026	646	274 (42.4)	F: 64.0 (7.1) M: 63.9 (7.5)	F: 31.5 (6.0) M: 29.8 (4.9)	F:61.2 (10.4) M: 60.4 (9.0)	1380	827 (59.9)	F: 58.3 (8.1) M: 58.7 (8.3)	F:26.1 (4.8) M: 26.2 (3.6)
European	The Dose Responses to Exercise Training (DR's EXTRA) Study	558	81	45 (55.5)	F: 66.8 (5.7) M: 68.2 (6.2)	F: 31.9 (5.4) M: 30.1 (4.6)	F: NA M: NA	477	360 (75.4)	F: 65.5 (5.1) M: 66.9 (6.3)	F: 26.7 (4.5) M: 26.3 (3.2)
European	National FINRISK 2007 Study (FINRISK 2007)	2606	1112	443 (39.8)	F: 61.4 (8.3) M: 59.4 (9.4)	F: 31.9 (6.1) M: 30.3 (4.3)	F: 58.3 (10.2) M: 56.8 (10.0)	1494	682 (45.6)	F: 62.5 (7.0) M: 60.0 (7.9)	F: 28.3 (4.8) M: 27.3 (3.4)
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	1467	981	439 (44.7)	F: 63.8 (7.9) M: 62.0 (7.3)	F: 31.0 (5.3) M: 29.3 (4.4)	F: 54.9 (8.7) M: 53.6 (8.0)	486	270 (55.5)	F: 62.6 (8.2) M: 62.3 (9.4)	F: 26.4 (4.0) M: 26.1 (3.5)
European	Prevalence, Prediction and Prevention of diabetes (PPP) (Finnish)	4969	311	127 (41)	F: 64.0 (10.8) M: 61.4 (12.5)	F: 30.1 (5.5) M: 30.1 (5.0)	F: 55.8 (15.4) M: 52.6 (12.9)	4658	2505 (54)	F: 48.8 (15.6) M: 48.6 (15.5)	F: 25.9 (4.7) M: 26.7 (3.8)
European	DIREVA (Diabetes Registry Vaasa) (Finnish)	2601	2601	1147 (44)	F: 67.3 (10.8) M: 65.8 (10.1)	F: 31.3 (5.6) M: 30.2 (5.2)	F: 59.9 (10.6) M: 58.6 (10.3)	0	F: NA M: NA	F: NA M: NA	F: NA M: NA
European	Malmö Diet and Cancer (MDC) (Swedish)	5613	440	192 (43)	F: 60.3 (5.3) M: 58.7 (5.8)	F: 30.4 (5.0) M: 29.4 (4.6)	F: NA M: NA	5173	3080 (55)	F: 57.3 (5.9) M: 57.3 (6.0)	F: 26.3 (4.1) M: 26.9 (3.5)
European	All New Diabetics In Scania (ANDiS) (Swedish)	1928	1928	776 (40)	F: 62.9 (11.6) M: 61.6 (11.3)	F: 32.9 (5.7) M: 31.6 (4.9)	F: 62.7 (11.5) M: 61.4 (11.2)	0	F: NA M: NA	F: NA M: NA	F: NA M: NA
European	Scania Diabetes Registry (SDR) (Swedish)	3192	3192	1312 (41)	F: 62.6 (13.2)	F: 30.5 (6.2)	F: 56.9 (14.0)	0	F: NA	F: NA	F: NA
European	Nurses' Health Study (NHS)	3088	1334	100 (100)	F: 43.4 (6.7) M: NA	F: 27.3 (4.9) M: NA	F: NA M: NA	1754	100 (100)	F: 43.2 (6.7) M: NA	F: 23.9 (3.0) M: NA
European	Health Professional Follow-Up Study (HPFS)	2411	1113	0 (0)	F: NA M: 55.5 (8.5)	F: NA M: 27.8 (4.0)	F: NA M: NA	1298	0 (0)	F: NA M: 55.5 (8.4)	F: NA M: 25.0 (2.7)
European	Estonian Genome Centre, University of Tartu (EGCUT)	2388	882	385 (43.7)	F: 62.0 (11.2) M:64.7 (10.8)	F: 31.6 (5.0) M:32.6 (5.6)	NA NA	1506	666 (44.2)	F: 47.2 (16.8) M: 46.8 (17.2)	F: 27.0 (4.1) M: 26.4 (5.3)
European	EFSOCH and DARE	3013	1446	564 (39.1)	F: 66.2 (9.02) M:65.9 (8.9)	F: 32.1 (6.2) M: 30.6 (5.2)	F:N/A M: N/A	1567	815 (52.3)	F: 30.4 (5.3) M: 32.9 (5.9)	F: 28.01 (4.6) M: 26.6 (3.9)
European	Cooperative Health Research in the Region of Augsburg [KORA]	3738	959	434 (45.3)	F: 62.7 (8.6) M: 61.2 (8.7)	F: 32.1 (5.7) M: 30.1 (4.7)	F: NA M: NA	2779	1436 (51.7)	F: 48.0 (13.1) M: 48.8 (13.2)	F: 26.5 (5.0) M: 27.3 (3.7)
European	Danish T2D case-control	10860	5864	2343 (40.0)	F: 61.2 (9.6) M: 61.4 (8.8)	F: 31.2 (6.2) M: 30.2 (4.9)	F: 54.9 (10.7) M: 54.2 (10.0)	4996	2716 (54.4)	F: 45.3 (7.9) M: 45.3 (7.9)	F: 25.0 (4.4) M: 26.1 (3.6)
European	GLACIER	1925	960	457 (47.60%)	F: 55.2 (6.7) M: 54.3 (7.5)	F: 30.6 (5.4) M: 29.5 (4.2)	NA NA	965	526 (54.51%)	F: 50.5 (7.7) M: 49.6 (8.5)	F: 25.4 (5.5) M: 26.1 (4.4)
European	EPIC-Norfolk (T2D cases) and the Fenland study (cohort)	1848	691	324 (47%)	F: 52.7 (9.9) M: 53.9 (10.1)	F: 27.6 (5.7) M: 28.1 (4.0)	F: 68.5 (8.4) M: 68.3 (8.3)	1157	631 (54.5%)	F: 48.5 (7.3) M: 48.2 (7.3)	F: 26.5 (5.6) M: 27.3 (4.0)

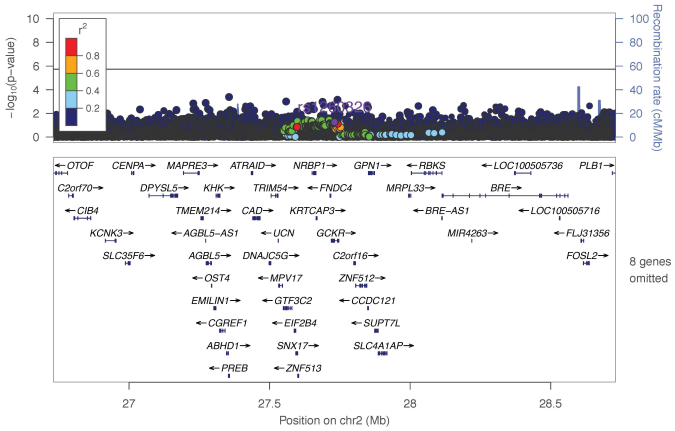
Supplementary Figure 16 | **Overlap of variants detected in 12,940 trans-ethnic exomes and genotyped on exome array in 79,854 Europeans.** Each blue bar indicates the number of coding SNVs, protein-altering SNVs (nonsense, essential splice site, and missense variants), or MAF > 0.5% protein-altering variants observed in 12,940 sequenced samples, broken down by ancestry (African American N=2,074; East Asian N=2,165; European N=4,541; Hispanic N=1,943; South Asian N=2,217). Red bars indicate the numbers of sequence variants that were observed in 79,854 European exome array samples. Exact counts are shown in the table on the right. While a small fraction of all coding variants are represented on exome array, 81.6% of European protein-altering variants with MAF >0.5% are captured using the array.



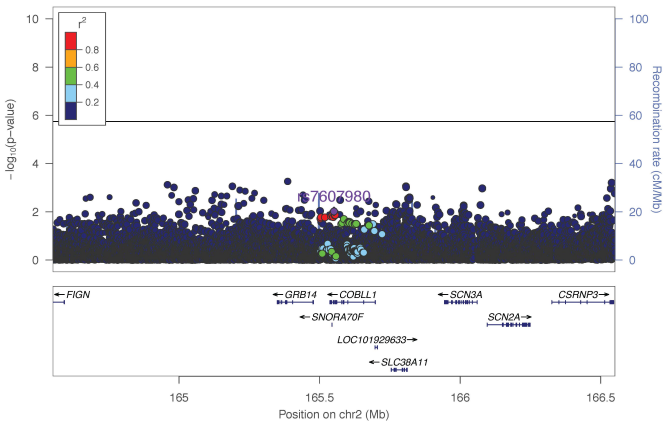
	# WES variants	# var on ex. chip	(%)
All coding SNVs			
All	1,738,367	94,792	5.5
Afr. Amer.	585,749	57,838	9.9
E. Asian	460,886	25,114	5.4
European	505,698	82,189	16.3
Hispanic	378,497	55,763	14.7
S. Asian	546,187	43,856	8.0
All protein-altering SNVs			
All	1,107,479	94,095	8.5
Afr. Amer.	349,722	57,362	16.4
E. Asian	288,193	24,873	8.6
European	318,730	81,589	25.6
Hispanic	225,259	55,326	24.6
S. Asian	338,134	43,484	12.9
Protein-altering SNVs with MAF > 0.5%			
All	79,097	37,573	47.5
Afr. Amer.	47,770	23,405	49.0
E. Asian	22,921	13,195	57.6
European	27,904	22,780	81.6
Hispanic	29,002	21,640	74.6
S. Asian	27,436	18,035	65.7

Supplementary Figure 17 | **Unconditional regional association plots for coding variants from GoT2D consortium data (N=2,657).** Each plot shows the p-value (on a -$\log_{10}$ scale) as a function of genomic position (NCBI Build 37) covering a 2-Mb window around the novel exome-wide significant coding variant (indicated by the purple symbol). The color-coding of all other SNPs indicates LD with the novel coding SNP estimated from GoT2D data: red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; green, 0.4 ≤ r2 < 0.6; cyan, 0.2 ≤ r2 < 0.4; blue, r2 < 0.2; gray, r2 unknown. Recombination rates are estimated from Phase II HapMap, and gene annotations are taken from the UCSC genome browser. Imputation quality was modest for rs60980157 (*GSPM1*) and rs9379084 (*RREB1*) (both r^2=0.84) and high for all other novel coding SNPs (r^2>0.99).

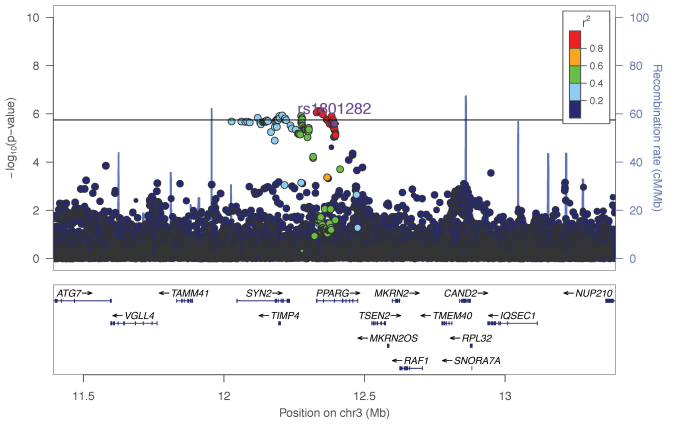
GCKR (rs1260326)



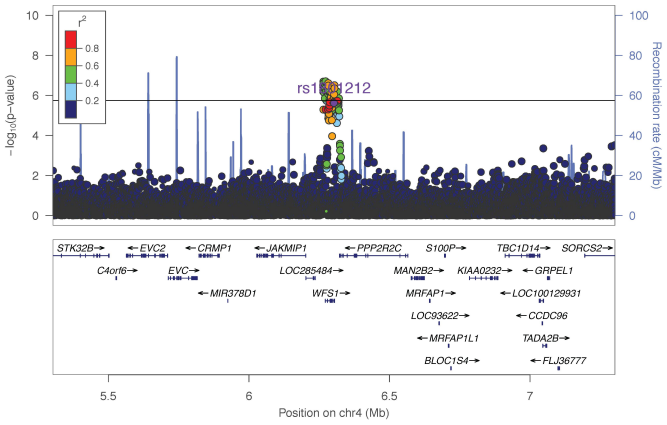
*COBLL*1 (rs7607980)



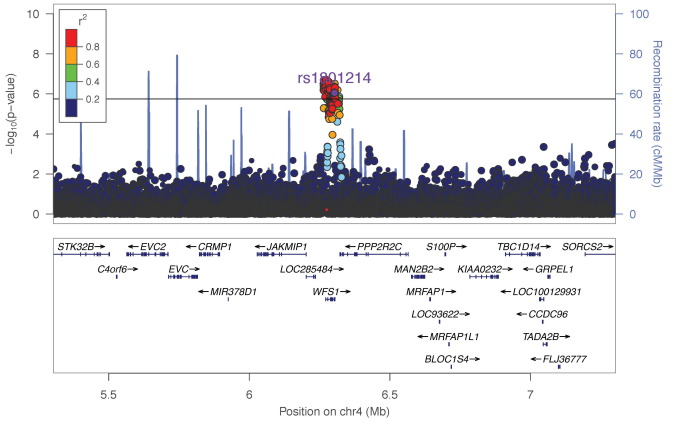
PPARG (rs1801282)



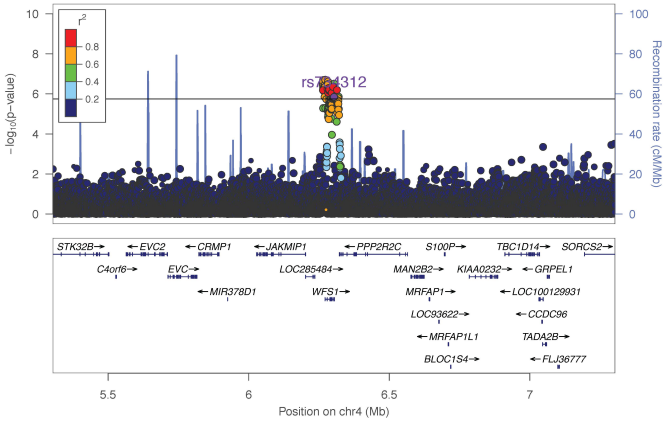
*WFS*1 (rs1801212)



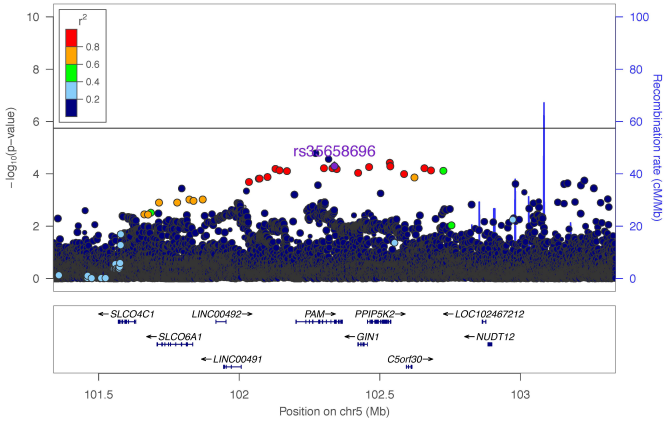
*WFS*1 (rs1801214)



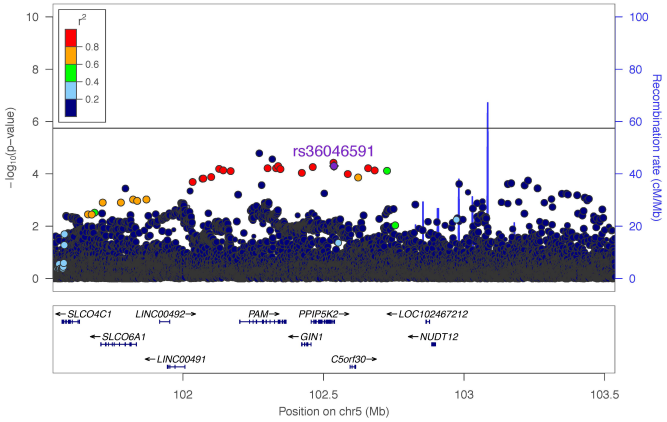
*WFS*1 (rs734312)



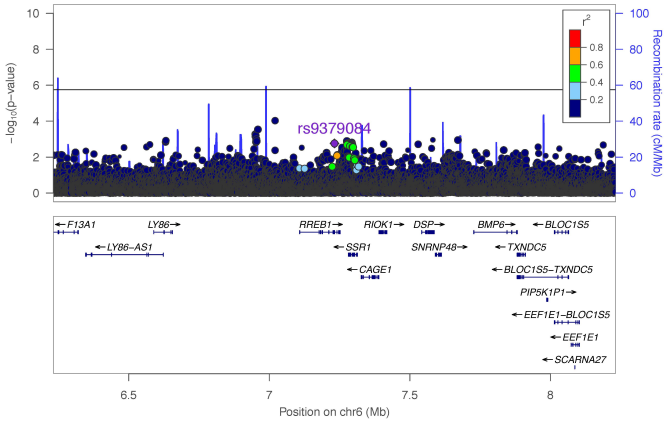
PAM (rs35658696)



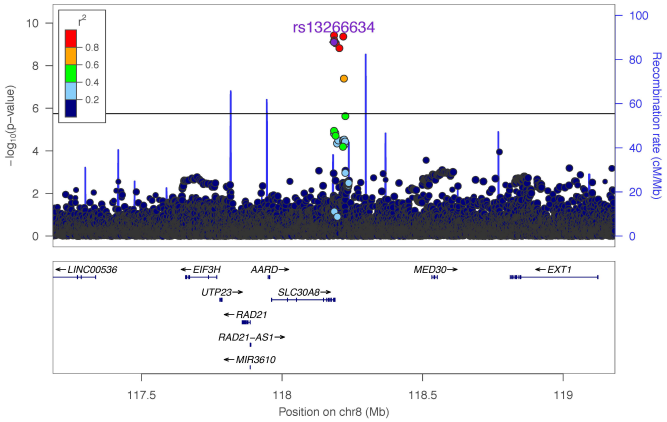
PPIP5K2 (rs36046591)



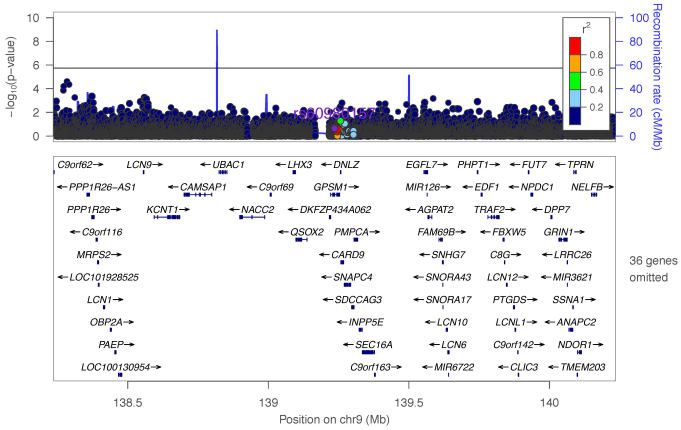
RREB1 (rs9379084)



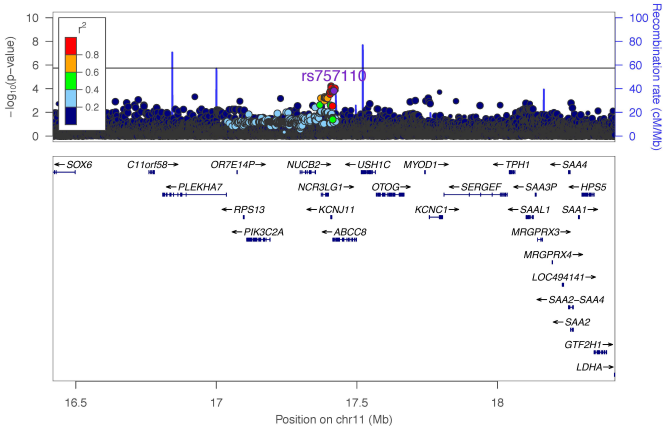
SLC30A8 (rs13266634)



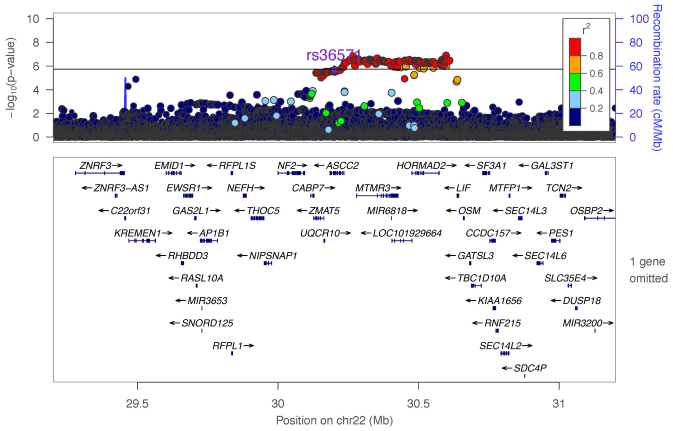
GPSM1 (rs60980157)



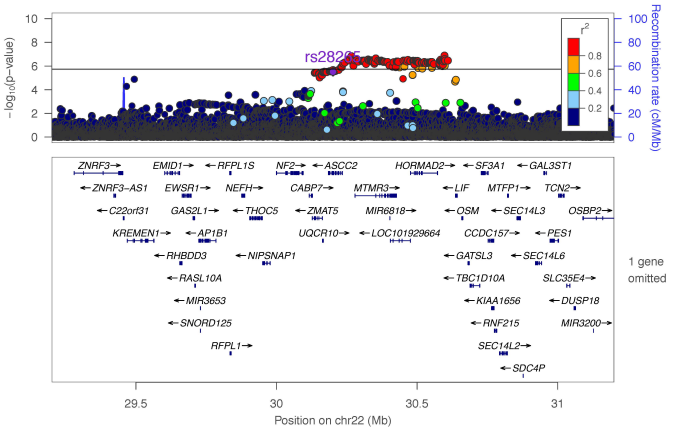
ABCC8 (rs757110)



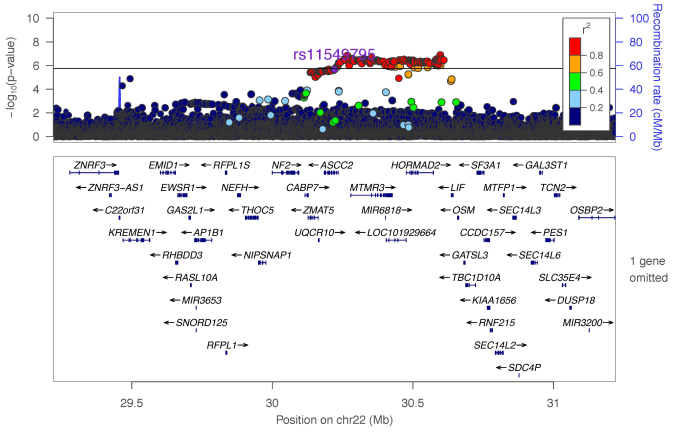
ASCC2 (rs36571)



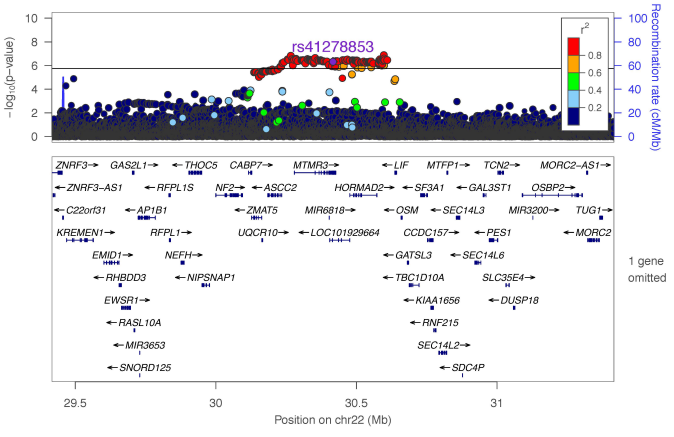
ASCC2 (rs28265)



ASCC2 (rs11549795)



MTMR3 (rs41278853)



Supplementary Table 18 | Association summary statistics for coding variants in GWAS regions from previous European ancestry meta-analysis from the DIAGRAM consortium.

Locus	rsID	Change	Combined exomes		Exome-array		GWAS (Morris et al. 2012)		GWAS + Metabochip (Morris et al. 2012)		Comments
			p-value	N	*p*-value	N	*p*-value	N	*p*-value	N	
PAM/PPIP5K2	rs35658696	Asp563Gly	$5.7x10^{-10}$	67,171	$1.7x10^{-7}$	66,047	NA	NA	NA	NA	Lead coding SNP from exome-array meta-analysis
	rs36046591	Ser1207Gly	$3.3x10^{-8}$	65,569	$1.0x10^{-6}$	62,201	NA	NA	NA	NA	
	rs7729395	non-coding	NA	NA	NA	NA	0.008	63,390	1.0×10^{-5}	80,645	Best proxy for the lead coding variant in Morris et al. (2012) (r^2=0.79)
MTMR3/ASCC2	rs41278853	Asn960Ser	$5.6x10^{-9}$	82,784	$3.2x10^{-6}$	79,852	NA	NA	NA	NA	Lead coding SNP from exome-array meta-analysis
	rs11549795	Val123Ile	$1.0x10^{-7}$	82,784	$2.0x10^{-5}$	79,854	0.0025	32,933	NA	NA	
	rs28265	Asp407His	$1.1x10^{-7}$	82,784	$1.9x10^{-5}$	79,849	0.0004	63,390	NA	NA	
	rs36571	Pro423Ser	$3.0x10^{-7}$	82,784	$2.0x10^{-5}$	79,854	0.00041	63,390	NA	NA	
	rs16988333	non-coding	NA	NA	NA	NA	0.00012	65,812	$3.7x10^{-6}$	82,788	Best proxy for the lead coding variant in Morris et al. (2012) (r^2=1)
	rs5997539	non-coding	NA	NA	NA	NA	$7.7x10^{-5}$	69,033	$7.7x10^{-6}$	80,480	Lead SNP from Morris et al. (2012) (r^2=0.74 with lead coding variant)
WFS1	rs1801214	Asn500Asn	$1.5x10^{-14}$	82,784	$2.0x10^{-12}$	79,854	$1.3x10^{-8}$	63,390	$3.3x10^{-15}$	80,640	Lead coding SNP from exome-array meta-analysis
	rs1801212	Val333Ile	$9.0x10^{-14}$	82,784	$9.3x10^{-12}$	79,852	$9.3x10^{-5}$	69,033	$3.6x10^{-11}$	83,539	
	rs734312	Arg611His	$6.9x10^{-11}$	82,783	$1.3x10^{-10}$	79,852	$3.2x10^{-7}$	63,390	$5.5x10^{-11}$	77,231	
	rs4689388	non-coding	NA	NA	$2.3x10^{-11}$	79,854	$3.3x10^{-7}$	63,390	$2.1x10^{-12}$	77,732	
	rs4458523	non-coding	NA	NA	NA	NA	$2.0x10^{-7}$	69,033	$2.0x10^{-15}$	85,051	Lead Morris et al. (2012) variant (r^2=1 with lead coding SNP from exome-array analysis)
CILP2/TM6SF2	rs58542926	Glu167Lys	$3.2x10^{-10}$	82,784	$1.9x10^{-7}$	79,854	NA	NA	$4.2x10^{-7}$	54,462	Lead coding SNP from exome-array meta-analysis
	rs10401969	non-coding	NA	NA	$4.2x10^{-7}$	79,854	0.00054	69,033	$7.0x10^{-9}$	86,196	
RREB1	rs9379084	Asp1171Asn	$4.0x10^{-9}$	56,339	$1.1x10^{-5}$	52,998	NA	NA	0.00022	54,618	Lead coding SNP from exome-array meta-analysis
	rs9502570	non-coding	NA	NA	NA	NA	0.00061	63,390	NA	NA	Lead SNP from Mahajan at el. (2014) (r^2=0.01 with the lead coding variant from exome-array analysis)
HNF4A	rs1800961	Thr139Ile	$2.9.x10^{-6}$	82,784	$9.5x10^{-7}$	79,854	0.025	60,203	0.00027	76,816	Lead coding SNP from exome-array meta-analysis
	rs4812831	non-coding	NA	NA	$4.7x10^{-5}$	79,854	0.016	69,033	NA	NA	Top SNP in Kooner et al. (2010) (r^2=0.01 with the lead coding variant from exome-array analysis)
THADA	rs35720761	Cys1650Tyr	$3.3x10^{-10}$	82,784	$3.5x10^{-8}$	79,845	NA	NA	NA	NA	Lead coding SNP from exome-array meta-analysis
	rs7578597	Thr1187Ala	$1.3x10^{-5}$	47,251	$5.1x10^{-5}$	37,704	$1.6x10^{-5}$	63,390	$2.0x10^{-9}$	78,010	
	rs10203174	non-coding	NA	NA	NA	NA	$1.5x10^{-6}$	69,033	$9.5x10^{-12}$	86,197	Lead Morris et al. (2012) variant (r^2=0.48 with lead coding SNP from exome-array analysis)
COBLL1	rs7607980	Asn939Asp	$8.3x10^{-15}$	82,784	$4.7x10^{-11}$	79,853	0.0067	69,033	$2.9x10^{-7}$	86,195	Lead coding SNP from exome-array meta-analysis
	rs13389219	non-coding	NA	NA	$1.9x10^{-10}$	79,850	0.0096	63,390	$1.0x10^{-8}$	80,649	Lead Morris et al. (2012) variant (r^2=0.77 with lead coding SNP from exome-array analysis)
TSPAN8	rs1051334	Ser213Ala	$2.7x10^{-6}$	62,197	$3.5x10^{-5}$	58,536	0.00018	63,390	0.00043	80,646	Lead coding SNP from exome-array meta-analysis
	rs4760790	non-coding	NA	NA	$8.6x10^{-6}$	79,854	$8.0x10^{-6}$	63,390	NA	NA	Lead SNP from Morris et al. (2012) (r^2=0.24 with lead coding variant)

Combined exomes p-values are derived from the meta-analysis of sequence and array datasets, with total sample size up to 92,794 (34,809 cases, 57,985 controls: effective sample size 82,758); smaller sample sizes reflect the fact that many variants were monomorphic in some or all of the non-European sequence cohorts. Exome-array analysis was performed in up to 79,854 samples (28,305 cases, 51,549 controls: effective sample size 69,866). Previously reported p-values from European meta-analysis are taken from Morris et al. (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 44(9):981-90.

Supplementary Table 19 | Association summary statistics for T2D and fasting glucose levels from exome-array and exome sequence for selected *RREB1* coding variants.

		T2D association							Fasting glucose (adjusted for BMI) association						
Asp1171Asn (rs9379084)															
		N_{eff}	**EAF (%)**	***p*-value**	**OR (95% CI)**	**I^2**	**Cochran's Q**	***p*_het**	**N**	**EAF (%)**	***p*-value**	**$\hat{\beta}$ (SE)**	**I^2**	**Cochran's Q**	***p*_het**
Within-ancestry	African Americans	2,074	97.6	0.00033	2.26 (1.81-2.71)	-	-	-	508	96.9	0.82	-0.040 (0.190)	-	-	-
	East Asians	2,165	86.6	0.062	1.20 (1.01-1.38)	-	-	-	1,104	85.7	0.22	0.075 (0.062)	-	-	-
	Europeans	4,541	89.5	0.0025	1.23 (1.09-1.37)	-	-	-	2,144	88.7	0.78	-0.014 (0.049)	-	-	-
	Hispanics	1,943	92.5	0.21	1.16 (0.92-1.41)	-	-	-	844	93.0	0.7	-0.037 (0.094)	-	-	-
	South Asians	2,217	89.9	0.071	0.86 (0.61-1.11)	-	-	-	508	87.2	0.39	0.085 (0.098)	-	-	-
Meta-analysis	Exome sequence	12,918	86.6-97.6	$2.2x10^{-5}$	1.19 (1.10-1.28)	67.2	12.2	0.016	5,108	85.7-96.9	0.57	0.019 (0.033)	0	2.2	0.70
	Exome-array	43,421	89	$1.1x10^{-5}$	1.12 (1.07-1.17)	2.8	9.3	0.41	24,031	88.6	0.0090	0.019 (0.007)	0	5.4	0.86
	Combined	56,339	86.6-97.6	$4.0x10^{-9}$	1.13 (1.09-1.18)	68.1	12.6	0.014	29,139	85.7-96.9	0.0084	0.034 (0.013)	0	2.5	0.78
Ser1554Tyr (rs35742417)															
Within-ancestry	African Americans	2,074	78.8	0.87	0.98 (0.83-1.13)	-	-	-	508	79.3	0.87	0.012 (0.076)	-	-	-
	East Asians	2,165	96.6	0.89	0.97 (0.64-1.31)	-	-	-	1,104	96.6	0.91	-0.013 (0.120)	-	-	-
	Europeans	4,541	82.3	0.023	1.12 (1.01-1.23)	-	-	-	2,144	80.9	0.28	0.042 (0.039)	-	-	-
	Hispanics	1,943	92.4	0.30	1.11 (0.87-1.35)	-	-	-	844	92	0.19	0.110 (0.087)	-	-	-
	South Asians	2,217	94.4	0.58	1.07 (0.81-1.33)	-	-	-	508	93.2	0.89	0.018 (0.120)	-	-	-
Meta-analysis	Exome sequence	12,918	78.8-96.6	0.065	0.94 (0.86-1.02)	0	3.2	0.53	5,108	80.9-96.6	0.17	0.041 (0.030)	0	1.1	0.90
	Exome-array	69,867	79.7	0.00029	1.05 (1.02-1.08)	21.4	15.3	0.23	33,230	78.9	$8.4x10^{-9}$	0.024 (0.004)	0	12.3	0.51
	Combined	82,784	78.8-96.6	$4.9x10^{-5}$	1.05 (1.02-1.08)	0	1.5	0.83	38,338	80.9-96.6	$2.7x10^{-9}$	0.053 (0.009)	0	1.3	0.94

N_{eff}: effective sample size. EAF: effect allele frequency. OR: odds-ratio. CI: confidence interval. I^2: heterogeneity measure in %. *p*_het: *p*-value for Cochran's Q statistic. N: number of individuals analysed. $\hat{\beta}$: regression coefficient estimates. SE: standard error.

Summary statistics of just the two coding variants showing significant association signals for either T2D or fasting glucose have been summarized.

For Supplementary Table 20 see Excel File “20Supp20 - T2D loci and genes.xlsx”

Supplementary Table 21 | Characterization of role of coding variants within genes in established common variant GWAS regions through reciprocal conditional analysis.

Locus	Variant	Combined exomes p-value	Coding variant(s)				GWAS variant				Comments
			Variant	Conditioned on	Unconditional *p*-value	Conditional *p*-value	Variant	Conditioned on	Unconditional *p*-value	Conditional *p*-value	
THADA	Cys1605Tyr	$3.3x10^{-10}$	Cys1605Tyr	unconditioned	0.00035		rs10203174	unconditioned	$5.7x10^{-6}$		The previously reported non-coding GWAS SNP (or a close proxy) at the *THADA* locus is not available on the exome array, so approximate conditional analyses were undertaken in GCTA in a genome-wide imputed meta-analysis from the GoT2D Consortium (**METHODS**). Association signals for *THADA* Cys1605Tyr and the GWAS SNP at this locus are partially correlated. The association signal for the GWAS SNP is not entirely extinguished in reciprocal conditional analysis. However, the association signals for *THADA* Thr1187Ala and the GWAS SNP at this locus are indistinguishable from each other in reciprocal conditional analysis. *THADA* is a candidate effector transcript for the non-coding GWAS signal at this locus.
				rs10203174		0.92		**Cys1605Tyr**		0.0063	
				Thr1187Ala		0.81					
	Thr1187Ala	$1.3x10^{-5}$	Thr1187Ala	unconditioned	$5.8x10^{-6}$		rs10203174	unconditioned	$5.7x10^{-6}$		
				rs10203174		0.37		**Thr1187Ala**		0.46	
				Cys1605Tyr		0.0053					
TSPAN8	Ser213Ala	$2.7x10^{-6}$	Ser213Ala	unconditioned	$3.5x10^{-5}$		rs4760790	unconditioned	$8.6x10^{-6}$		Association signals for *TSPAN8* Ser231Ala and the GWAS SNP at this locus are partially correlated. The association signal for the GWAS SNP is not entirely extinguished in reciprocal conditional analysis. *TSPAN8* is a candidate effector transcript at this locus.
				rs4760790		0.00024		**Ser213Ala**		0.0025	
HNF4A	Thr139Ile	$2.9.x10^{-6}$	Thr139Ile	unconditioned	$9.5x10^{-7}$		rs4812831	unconditioned	$4.7x10^{-5}$		Association signals for *HNF4A* Thr139Ile and the GWAS SNP at this locus are independent of each other. The association signal for the GWAS SNP is not extinguished in reciprocal conditional analysis. Previous GWAS signal is not mediated through *HNF4A* Thr139Ile.
				rs4812831		$4.0x10^{-7}$		**Thr139Ile**		$2.4x10^{-5}$	
				rs10842994		0.18		**Gly43Arg**		$7.7x10^{-5}$	

Combined exomes p-values are derived from the meta-analysis of sequence and array datasets, with total sample size up to 92,794 (34,809 cases, 57,985 controls: effective sample size 82,758). Conditional analysis was performed only on the exome-array component (28,305 cases, 51,549 controls: effective sample size 69,866). However, the previously reported non-coding GWAS SNP at the *THADA* locus (rs10203174) is not available on the exome array; p-values reported here come from approximate conditional analyses undertaken in GCTA in a genome-wide imputed meta-analysis from the GoT2D Consortium (**METHODS**). We also examined genome-wide sequence and imputed data sets from the GoT2D consortium (N=2,657; **METHODS**) to determine whether these causal inferences were robust to more comprehensive coverage of regional variation.

Supplementary Table 22 | List of monogenic analysis categories and genes

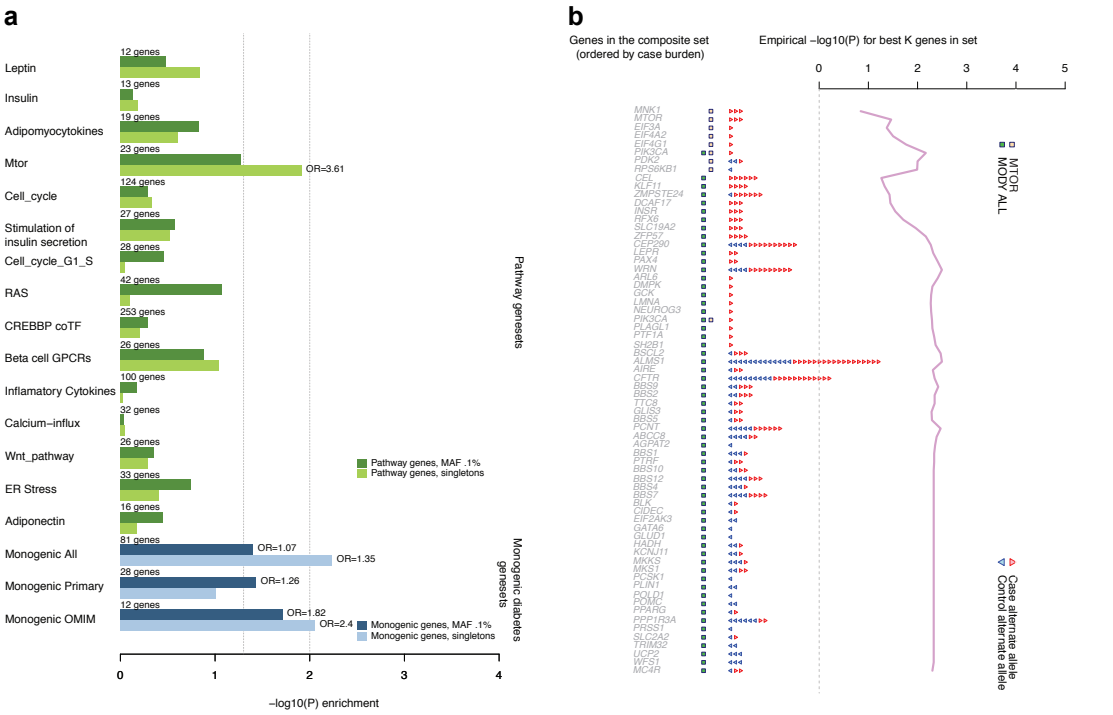
Monogenic ALL	Monogenic PRIMARY	Monogenic OMIM
ABCC8	*ABCC8*	*ABCC8*
AGPAT2		
AIRE		
AKT2		
ALMS1	*ALMS1*	
ARL6		
BBS1		
BBS10		
BBS12		
BBS2		
BBS4		
BBS5		
BBS7		
BBS9		
BLK		*BLK*
BSCL2		
CAV1		
CEL	*CEL*	*CEL*
CEP290		
CFTR		
CIDEC		
CISD2	*CISD2*	
DCAF17		
DMPK		
EIF2AK3	*EIF2AK3*	
FOXP3		
GATA4	*GATA4*	
GATA6	*GATA6*	
GCK	*GCK*	*GCK*
GLIS3	*GLIS3*	
GLUD1	*GLUD1*	
HADH	*HADH*	
HNF1A	*HNF1A*	*HNF1A*
HNF1B	*HNF1B*	*HNF1B*
HNF4A	*HNF4A*	*HNF4A*
INS	*INS*	*INS*
INSR	*INSR*	
ISL1		
ITER3IP1		
KCNJ11	*KCNJ11*	*KCNJ11*
KLF11		*KLF11*
LEP		
LEPR		
LMNA	*LMNA*	
MC4R		
MKKS		
MKS1		
MNX1	*MNX1*	
NEUROD1	*NEUROD1*	*NEUROD1*
NEUROG3	*NEUROG3*	
PAX4		*PAX4*
PAX6		
PCNT		
PCSK1		
PDX1	*PDX1*	*PDX1*
PIK3CA		
PIK3R1		
PLAGL1		
PLIN1		
POLD1		
POMC		
PPARG	*PPARG*	
PPP1R3A		
PRSS1		
PTEN		
PTF1A	*PTF1A*	
PTRF		
RFX6	*RFX6*	
SH2B1		
SIRT1		

SLC16A1		
SLC19A2		
SLC2A2	*SLC2A2*	
SPINK1		
TRIM32		
TTC8		
UCP2	*UCP2*	
WFS1	*WFS1*	
WRN		
ZFP57		
ZMPSTE24		

Supplementary Figure 23 | Age of diagnosis of variant carriers. To assess whether individuals carrying variants in genes for monogenic forms of diabetes were enriched for patients with undiagnosed monogenic diseases, we examined the ages of diagnosis for carriers and compared them to those of non-carriers. As some diseases typically manifest at an earlier age than does T2D (e.g. MODY), a lower age of diagnosis for carriers might suggest that the monogenic phenotype, rather than late-onset T2D, is responsible for the diabetes phenotype in carriers. Shown is the mean age of diagnosis for carriers of variants in (a) the Monogenic All gene set, (b) the Monogenic Primary gene set, and (c) the Monogenic OMIM gene set. In each case, the mean ages are computed for carriers of variants in each of the five variant masks discussed in the text. Error bars indicate one standard deviation. Numbers of carriers and non-carriers for each mask are listed in parentheses at the bottom of the plot.

a
Monogenic All
Mean Age of Diagnosis
0
20
40
Carriers (3299)
Non – carriers (740)
Carriers (2905)
Non – carriers (1134)
Carriers (1150)
Non – carriers (2889)
Carriers (253)
Non – carriers (3786)
Carriers (3755)
Non – carriers (284)
PTV + missense
PTV + NS$_{Broad}$
PTV + NS$_{Strict}$
PTV – only
HGMD
b
Monogenic Primary
Mean Age of Diagnosis
0
20
40
Carriers (1904)
Non – carriers (2135)
Carriers (1488)
Non – carriers (2551)
Carriers (384)
Non – carriers (3655)
Carriers (47)
Non – carriers (3992)
Carriers (653)
Non – carriers (3386)
PTV + missense
PTV + NS$_{Broad}$
PTV + NS$_{Strict}$
PTV – only
HGMD
c
Monogenic OMIM
Mean Age of Diagnosis
0
20
40
Carriers (849)
Non – carriers (3190)
Carriers (670)
Non – carriers (3369)
Carriers (201)
Non – carriers (3838)
Carriers (19)
Non – carriers (4020)
Carriers (514)
Non – carriers (3525)
PTV + missense
PTV + NS$_{Broad}$
PTV + NS$_{Strict}$
PTV – only
HGMD

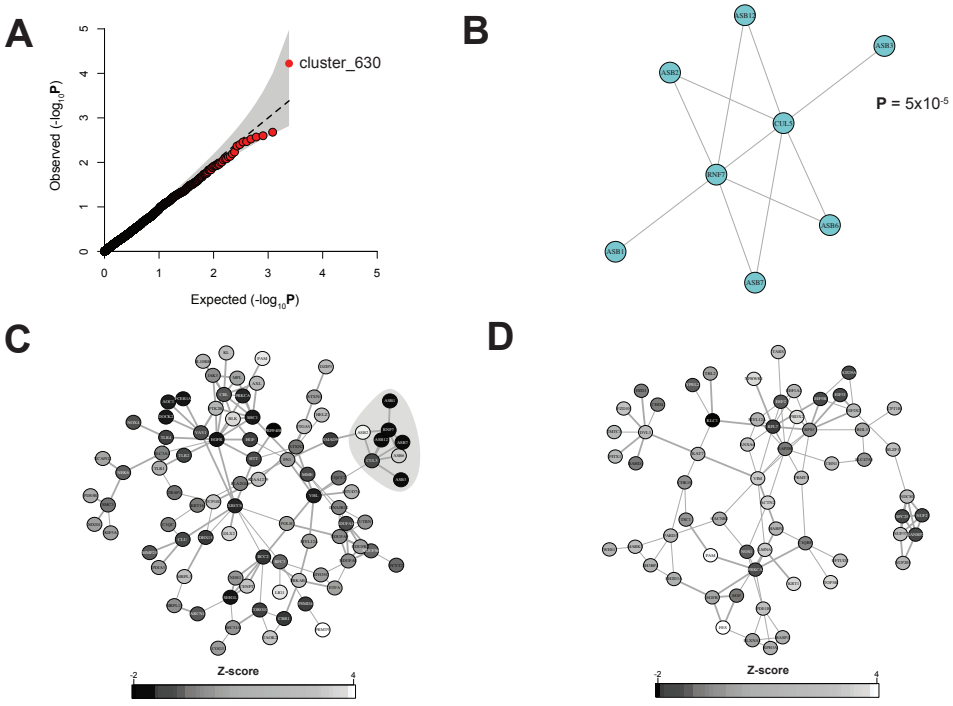
Supplementary Figure 24 | Accumulation of ultra-rare deleterious alleles amongst genes contributing to pre-specified "biologically-driven" gene-sets. a, Using the SMP approach, we confirmed association in the 'Monogenic All" gene set (81 genes: p= 0.006, OR=1.35 for singletons; p=0.04, OR=1.07 for ultra-rare alleles) and the "Monogenic OMIM" gene set (13 genes: p=0.0088, OR=2.4 for singletons; p=0.02, OR=1.82 for ultra-rare alleles). We also detected a separate "burden" signal for increased T2D-risk attributable to singleton alleles within the MTOR pathway (p=0.012, OR=3.61). **b**, Individual gene-ranking of composite set genes (set genes with *p*< .05 are shown). Genes are ranked by their case burden of rare PTVs, from top to bottom, for the Mtor and the monogenic all gene sets (labeled MTOR and MODY ALL, respectively). The squares along the bottom indicate to which sets each gene belongs. The red and blue triangles represent case and control counts for each gene. The lines represent the statistical significance of the best test for this set: that is, the significance of the top K genes, evaluated by permutation. For example, the drivers of the MTOR pathway signal include three case-only PTV singletons in both MNK1 and MTOR.



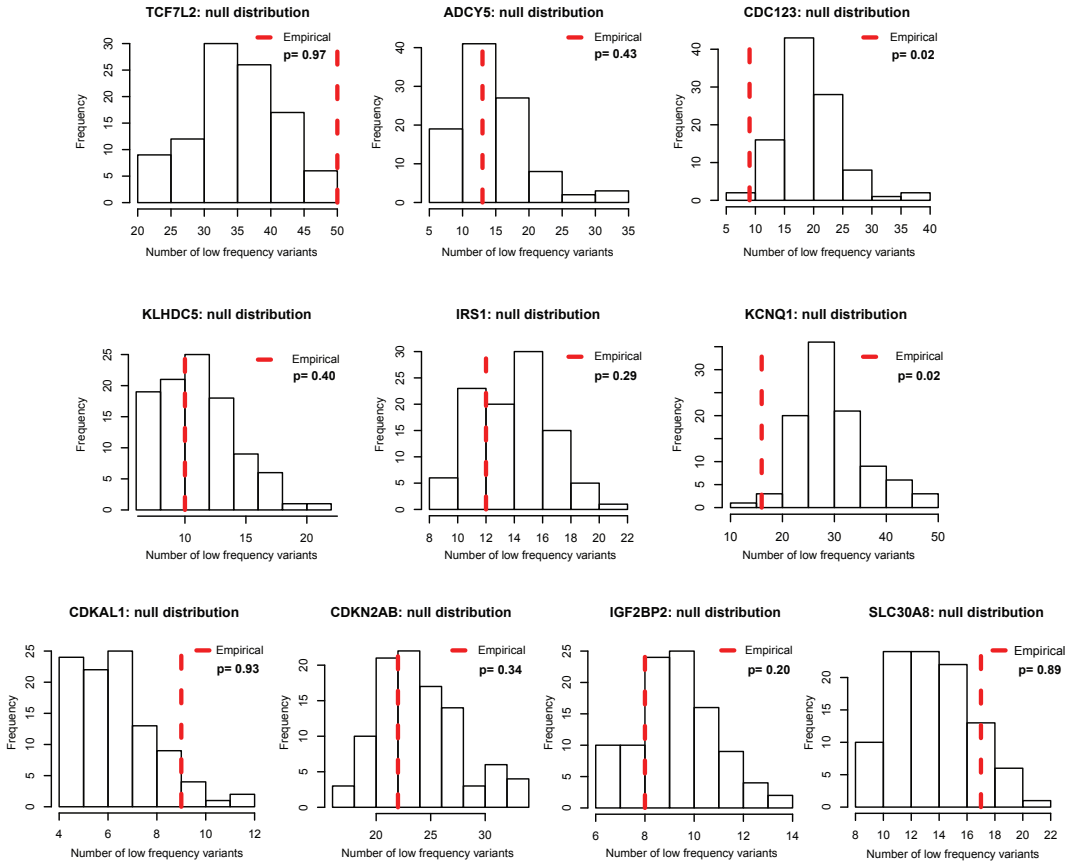
Supplementary Table 25 | All pathway enrichment signals with uncorrected FDR <= 10% from gene-set enrichment analyses conducted on the ancestry-specific and combined exome sequence data. Ancestries are denoted as European (EUR), East Asian (EA), South Asian (SA), Hispanic (HS) and African American (AA). For hand-curated gene-sets see **Supplementary Table 32**. We detected no study-wide significant signals (defined FDR < 5% after correction for multiple testing on four masks and five gene set collections). However, we detect nominal associations (uncorrected FDR <= 10%) in a subset of analyses, as listed below.

Mask	Ancestry	GeneSet	FDR
PTV+NS_{broad}	Combined	Transport of organic anions (Reactome)	1.1%
PTV+missense	AA	Cytosolic DNA sensing pathway (Kegg)	1.4%
PTV+NS_{strict}	HS	Glycosaminoglycan biosynthesis chondroitin sulphate (Kegg)	3.1%
PTV only	HS	TRIF-mediated TLR3 signaling (Reactome)	3.5%
PTV+NS_{strict}	SA	Secretory granule (Gene Ontology)	3.6%
PTV+NS_{broad}	AA	Riboflavin metabolism (Kegg)	4.9%
PTV+NS_{broad}	Combined	Mendelian: long qt syndrome (hand curated)	5.1%
PTV only	EA	Fatty acid metabolism (Kegg)	5.7%
PTV+NS_{strict}	EUR	Mismatch repair (Kegg)	5.7%
PTV+NS_{broad}	Combined	Mendelian: immune (hand curated)	5.8%
PTV+NS_{strict}	EA	Beta alanine metabolism (Kegg)	6.4%
PTV only	AA	Tak1 activates NFKB by phosphorylation and activation of IKKS complex (Reactome)	6.9%
PTV+NS_{broad}	HS	ABC transporters (Kegg)	7.5%
PTV+NS_{broad}	Combined	Cellular polysaccharide metabolic process (Gene Ontology)	8.4%
PTV+NS_{strict}	EA	Limonene and pinene degradation (Kegg)	8.6%
PTV+NS_{broad}	EA	Integrin signaling (MSigDb Canonical Pathway)	8.9%
PTV+NS_{strict}	AA	Triggering pathway mediating stimulation of insulin secretion (hand-curated)	9.5%
PTV+NS_{strict}	AA	Lysine degradation (Kegg)	10.0%

Supplementary Figure 26 | PPI analyses. **A,** QQ-plot of Fisher aggregated empirical *p*-values ("PTV+NS_{strict}" mask) from the 2418 clusters generated by clusterONE based on 100,000 iterations. Cluster 630, consisting of ASB (ankyrin repeat and SOC box protein) protein family members interacting with RNF7 and CUL5, showed the strongest enrichment ("PTV+NS_{strict}" mask, empirical *p*-value P=5x10^{-5}); **B**, Membership of the cluster 630 sub-network highlighted by the clusterONE analyses. ASB6 is adipocyte-specific and interacts with APS to enable recruitment of elongins to the insulin receptor-signaling complex; **C**, PPI sub-network constructed using the top 15 modules generated with dmGWAS from gene-based association *p*-values derived using the "PTV+NS_{strict}" variant mask. The sub-network includes the cluster of ASB proteins found in the clusterONE method as significant (cluster 630, shaded area), a cluster of mitochondrial-activity related genes, and the *PAM* gene; **D,** PPI sub-network built using the 15 top modules generated with dmGWAS from gene-based association *p*-values derived using the "PTV+NS_{broad}" mask. The sub-network includes *PAM* and *FES*, both of which contain exome-wide significant coding variants associated with T2D. The darkness of the node in the sub-networks is proportional to its *p*-value (lighter color indicates lower *p*-values) and the thickness of the edge is proportional to confidence score for the interaction between each pair of proteins.



Supplementary Figure 27 | Use of permutations to evaluate synthetic association hypothesis at 10 T2D GWAS loci.



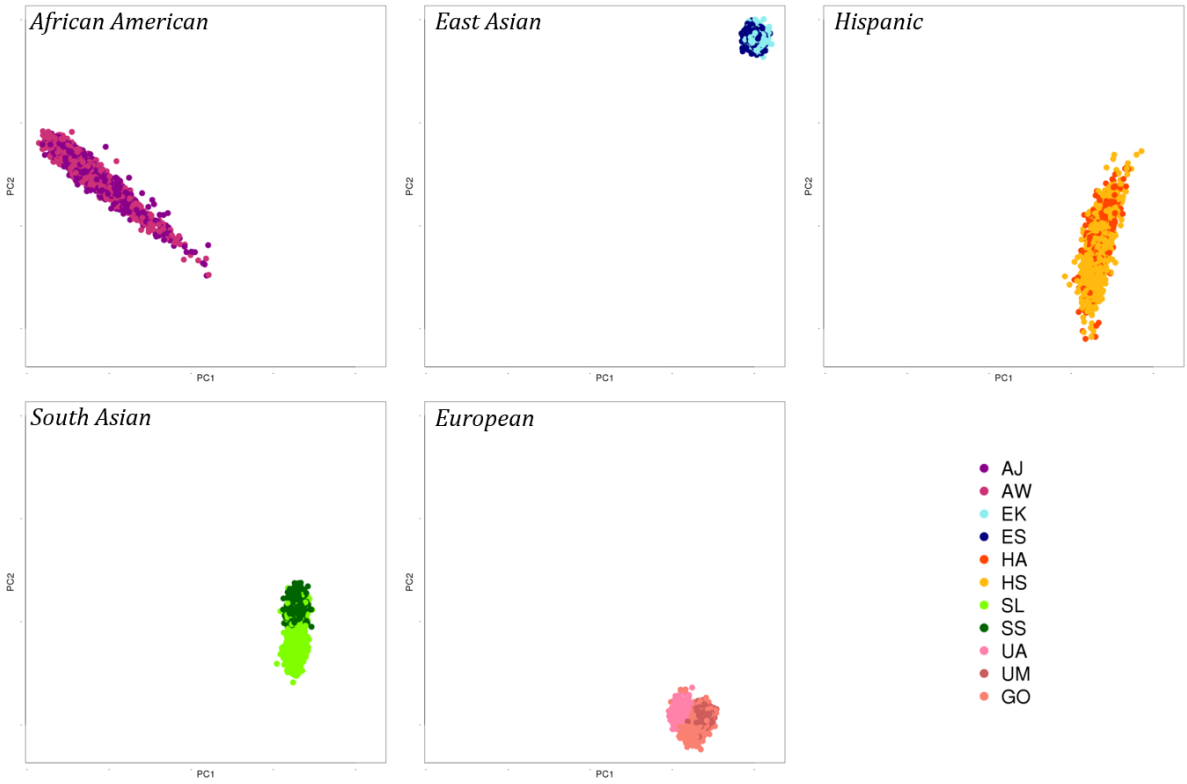
Supplementary Table 28 | Properties of credible sets constructed at all T2D GWAS loci. Loci are sorted by the size of the final 99% credible set (from smallest to largest; column 5 below). Up to 2 credible sets were constructed for independent signals (r^2<0.1) at all previously known autosomal T2D GWAS loci. Only loci where the index SNV had MAF>1% in the GoT2D sequencing data were included; *RBM43* and *SGCG* were excluded due to low index SNV MAF. *CILP2* was excluded due to poor sequencing quality across this region in the GoT2D experiment. At loci where the two independent signals have opposite directions of effect at the minor allele (risk, protective), the credible sets are labeled as such ("risk", "prot"); at loci where both signals are in the same direction, they are labeled "sig1" and "sig2".

Locus	# candidate variants with r^2>0.1 to GWAS tag	# indels	# imputed with low quality	# variants in 99% credible set	# missense variants in credible set	% variants in 1000G	% variants in HapMap	Post. prob. of top 5 variants	Post. prob. of top 10 variants
CDKN2AB_risk	16	0	1	2	-	100%	50%	1.00	1.00
TCF7L2	121	8	0	3	-	100%	33%	1.00	1.00
CCND2	27	4	1	4	-	100%	25%	0.99	0.99
ZBED3	99	2	0	5	-	100%	60%	0.99	0.99
KCNQ1_sig2	59	4	0	5	-	100%	40%	1.00	1.00
SLC30A8	57	3	0	7	1	100%	43%	0.93	0.99
CDKAL1	376	33	1	9	-	100%	56%	0.74	1.00
CDKN2AB_prot	27	0	2	12	-	100%	42%	0.75	0.96
HHEX	236	23	2	14	-	100%	43%	0.47	0.79
BCL11A	189	16	0	15	-	100%	73%	0.76	0.96
ST6GAL1	46	1	1	19	-	100%	42%	0.66	0.92
HNF1B	40	2	4	21	-	100%	48%	0.80	0.92
ADCY5	200	20	3	22	-	100%	41%	0.71	0.86
JAZF1	171	13	2	25	-	88%	40%	0.52	0.80
ADAMTS9	133	8	0	26	-	100%	62%	0.49	0.79
TLE1	185	18	1	26	-	100%	42%	0.86	0.94
GCK	29	2	0	28	-	100%	57%	0.49	0.61
DUSP8	110	10	4	28	-	93%	36%	0.58	0.81
PROX1	88	7	0	29	-	93%	59%	0.57	0.82
BCAR1	616	42	13	30	-	90%	10%	0.66	0.83
ZMIZ1	164	8	2	33	-	100%	70%	0.54	0.82
BCL2	35	1	0	35	-	97%	69%	0.56	0.64
KCNQ1_sig1	83	2	35	40	-	98%	18%	0.72	0.76
TSPAN8	201	15	2	41	-	95%	54%	0.73	0.94
SPRY2	206	14	0	42	-	95%	43%	0.54	0.95
LAMA1	56	3	3	45	-	96%	62%	0.25	0.44
ANK1	176	17	0	48	-	96%	35%	0.37	0.65
IGF2BP2	121	11	1	50	-	96%	40%	0.17	0.29

CENTD2	311	23	4	53	-	94%	23%	0.38	0.71
PTPRD	55	5	1	54	-	96%	35%	0.26	0.37
PRC1	260	27	1	62	1	94%	45%	0.32	0.50
HMGA2	209	13	3	67	-	94%	48%	0.38	0.68
LPP	121	11	0	70	-	96%	49%	0.71	0.79
SLC16A13	78	5	10	71	5	99%	30%	0.15	0.28
KLHDC5	212	27	0	72	-	94%	26%	0.28	0.48
CDC123	170	16	7	74	-	99%	32%	0.64	0.88
FTO	144	4	1	80	-	100%	41%	0.30	0.48
MTNR1B	91	7	1	81	-	93%	37%	0.48	0.56
UBE2E2	507	37	1	84	-	98%	30%	0.37	0.68
IRS1	403	37	0	90	-	99%	49%	0.16	0.28
ARL15	245	25	0	92	-	93%	28%	0.10	0.18
KLF14	98	8	1	93	-	98%	34%	0.09	0.17
RASGRP1	96	6	2	93	-	100%	45%	0.22	0.32
DGKB_prot	154	8	5	101	-	98%	53%	0.17	0.31
C2CD4	229	9	0	102	-	99%	60%	0.18	0.31
MC4R_prot	146	12	17	107	1	98%	14%	0.12	0.23
PAX4_prot	115	6	2	111	-	97%	31%	0.14	0.23
TMEM154	135	10	8	113	-	96%	53%	0.28	0.39
KCNK16	119	11	0	117	3	99%	49%	0.11	0.20
PPARG	306	28	1	123	1	95%	33%	0.12	0.20
NOTCH2	137	11	3	128	2	96%	52%	0.24	0.30
HNF4A	202	18	0	139	-	96%	34%	0.74	0.85
GLIS3	145	6	0	139	-	98%	57%	0.19	0.30
ZFAND6	165	15	0	144	-	97%	33%	0.24	0.36
WFS1	210	16	1	145	2	93%	39%	0.24	0.36
GRK5	148	14	0	146	-	97%	51%	0.18	0.27
ANKRD55	181	7	0	148	-	100%	31%	0.27	0.47
VPS26A	152	12	0	150	-	97%	35%	0.07	0.12
GIPR	238	18	1	151	2	96%	21%	0.82	0.83
AP3S2	186	18	8	161	-	98%	39%	0.15	0.25
DGKB_risk	199	13	3	166	-	97%	46%	0.51	0.73
GRB14	177	17	1	170	1	99%	51%	0.19	0.34
PEPD	316	30	0	194	-	98%	30%	0.43	0.72
GPSM1	223	17	26	220	5	96%	40%	0.06	0.09

TLE4	326	21	13	224	-	99%	50%	0.19	0.32
MC4R_risk	388	33	0	250	-	97%	50%	0.10	0.19
RBMS1	313	16	0	250	-	97%	45%	0.23	0.31
SSR1	313	38	0	256	2	97%	22%	0.10	0.19
CTBP1	279	20	23	263	1	95%	13%	0.15	0.26
MACF1	429	37	8	305	1	94%	30%	0.13	0.18
HMG20A	955	68	0	319	-	97%	44%	0.10	0.17
THADA	592	41	5	319	3	97%	39%	0.07	0.12
HNF1A	516	45	1	366	3	95%	35%	0.42	0.56
KCNJ11	528	50	1	398	4	96%	36%	0.52	0.75
MPHOSPH9	477	39	57	437	1	96%	32%	0.03	0.05
ZFAND3	466	44	6	447	-	99%	25%	0.09	0.12
GCKR	468	51	1	451	8	96%	43%	0.24	0.27
SRR	463	34	0	457	5	98%	43%	0.11	0.14
TP53INP1	696	53	0	502	-	97%	36%	0.16	0.24
FAF1	839	75	10	773	1	97%	18%	0.19	0.20
TMEM163	887	83	37	835	1	96%	43%	0.20	0.28
PAX4_risk	857	87	0	844	1	96%	49%	0.03	0.05
POU5F1	1782	74	1	993	14	96%	37%	0.24	0.30
					Averages	97%	41%		

Supplementary Figure 29 | Trans-ethnic principal component analysis for exome-sequence samples. African American studies (N=2,074): Jackson Heart Study (AJ) and Wake Forest School of Medicine Study (AW); East Asian studies (N=2,165): Korea Association Research Project (EK) and Singapore Diabetes Cohort Study and Singapore Prospective Study Program (ES); Hispanic studies (N=1,943): San Antonio Family Heart Study (HA) and Starr County (HS); South Asian studies (N=2,217): London Life Sciences Population Study (SL) and Singapore Indian Eye study (SS); and European studies (N=4,541): Ashkenazi (UA), Metabolic Syndrome in Men Study (UM), and GoT2D study (GO). A total of 10,348 independent QC passed, autosomal variants (trans-ethnic r^2<0.05) with MAF>1% in all ancestry groups were considered for constructing axes of genetic variation through principal components analysis implemented in EIGENSTRAT to identify ethnic outliers.

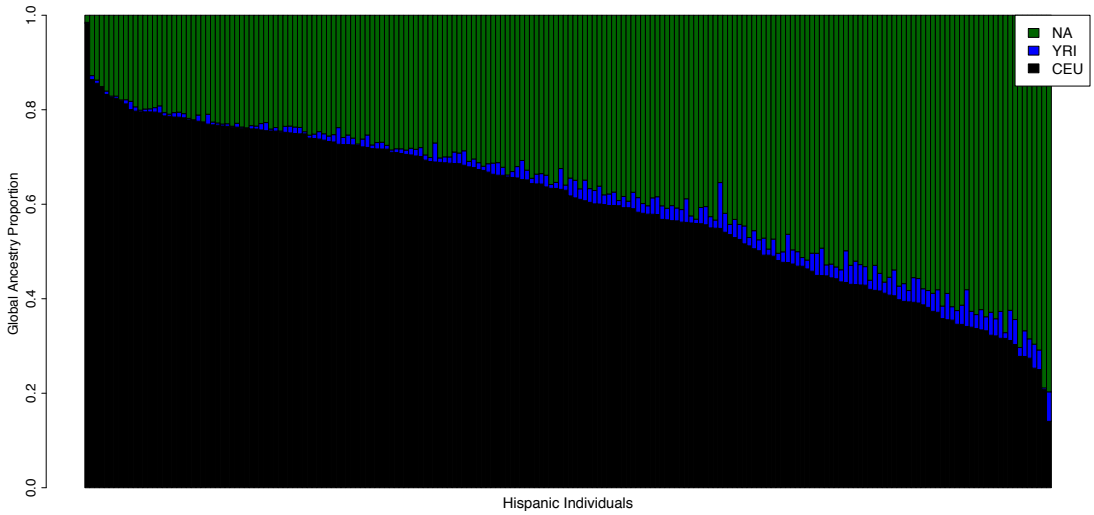


Supplementary Table 30 | **Summary of samples and adjustments for EMMAX and WALD single-variant association analysis for 12,940 exome-sequence samples**. Genomic control inflation factors (λ) were calculated on the basis of independent autosomal variants with MAF>1% within each ancestry group.

Ancestry group	EMMAX		WALD		
	N	λ	N	PCs used for adjustment	λ
African American	2,074	1.00	2,025	PC1	1.01
East Asian	2,165	1.02	2,164	PC1	1.02
European	4,541	0.99	4,518	PC1 - PC4	1.07
Hispanic	1,943	1.00	1,938	PC1 - PC4	1.03
South Asian	2,217	1.03	2,199	PC1	1.05

N: total number of samples. PC: principal component.

Supplementary Figure 31 | Global ancestry estimates for 1,943 Hispanic samples.



Supplementary Table 32 | Premium gene sets

Diabetes_McGill_Adipomyocytokines	Diabetes_McGill_Adipomyocytokines	*ADIPOQ AGT APLN AZGP1 C3 FGF21 IL10 IL13 IL1B IL6 IL8 LEP CCL2 MIF NAMPT RETN TGFB1 TNF VEGFA*
Diabetes_McGill_Adiponectin_Pathway	Diabetes_McGill_Adiponectin_Pathway	*ACACA ACACB ADIPOQ ADIPOR1 ADIPOR2 APPL1 SLC2A4 PRKAA1 PRKAA2 PRKAB1 PRKAB2 PRKAG1 PRKAG2 PRKAG3 RAB5A STK11*
Diabetes_McGill_Beta_cell_GPCRs	Diabetes_McGill_Beta_cell_GPCRs	*GPR119 ADRB2 ADRA2A MTNR1A MTNR1B HTR2B HTR1D CHRM3 SSTR2 GLP1R FFAR1 GCGR ADRB1 GIPR FFAR3 FFAR2 O3FAR1 ADCYAP1R1 VIPR2 NPY1R GHSR KISS1R CCKAR CNR1 P2RY12 ADRA2B*
Diabetes_McGill_Calcium-influx_pathway_Katp-independent_and_dependent	Diabetes_McGill_Calcium-influx_pathway_Katp-independent_and_dependent	*ABCC8 KCNJ11 SSTR2 SSTR3 KCNJ3 KCNJ6 NALCN CACNA1H CACNA1D CACNA1A CACNA1C CACNA1E SCN8A SCN9A SCN1B SCN3B KCNMA1 KCNB1 KCNB2 KCNQ1 KCNH2 KCNJ12 KCNJ4 KCNJ15 KCNN3 KCNN4 KCNN1 RYR2 ITPR1 ITPR2 ITPR3 CLCN3*
Diabetes_McGill_Cell_cycle_All_Genes	Diabetes_McGill_Cell_cycle_All_Genes	*CCND1 CCND2 CCND3 CDK4 CDK6 RB1 RBL1 RBL2 ABL1 HDAC1 HDAC2 E2F1 E2F2 E2F3 E2F4 E2F5 TFDP1 TFDP2 GSK3B TGFB1 TGFB2 TGFB3 SMAD2 SMAD3 SMAD4 MYC ZBTB17 CDKN2A CDKN2B CDKN2C CDKN2D CDKN1B CDKN1C CDKN1A CCNE1 CCNE2 CDK2 SKP1 CUL1 RBX1 SKP2 CCNA2 CCNA1 CDC6 CDC45 CDC7 DBF4 CDK1 CCNB1 CCNB2 CCNB3 CDC25B CDC25C YWHAZ YWHAB YWHAQ YWHAE YWHAH YWHAG PLK1 WEE1 WEE2 PKMYT1 CCNH CDK7 ANAPC1 ANAPC2 CDC27 ANAPC4 ANAPC5 CDC16 ANAPC7 CDC23 ANAPC10 ANAPC11 CDC26 ANAPC13 CDC20 PTTG1 PTTG2 ESPL1 SMC1A SMC1B SMC3 STAG2 STAG1 RAD21 TTK BUB1 BUB3 BUB1B MAD1L1 MAD2L1 MAD2L2 FZR1 CDC14B CDC14A ATR ATM TP53 CHEK1 CHEK2 CREBBP EP300 PRKDC MDM2 GADD45A GADD45B GADD45G PCNA SFN CDC25A ORC1 ORC2 ORC3 ORC4 ORC5 ORC6 MCM2 MCM3 MCM4 MCM5 MCM6 MCM7*
Diabetes_McGill_Cell_cycle_G1_S	Diabetes_McGill_Cell_cycle_G1_S	*ABL1 ATM ATR CCNA1 CCND1 CCNE1 CDC25A CDK1 CDK2 CDK4 CDK6 CDKN1A CDKN1B CDKN2A CDKN2B DHFR E2F1 GSK3B HDAC1 RB1 SKP2 SMAD3 SMAD4 TFDP1 TGFB1 TGFB2 TGFB3 TP53*
Diabetes_McGill_CREBBP_coTF	Diabetes_McGill_CREBBP_coTF	*SREBF1 IRF9 KLF5 SERTAD1 AC091153.1 NCOA2 TDG FHL2 ZBTB2 RBBP4 ONECUT1 PLAGL1 VDR HDAC3 EBF1 ACVR1 CSNK2A1P HOXB4 NCOA6 NCOA3 SPIB HOXB7 RELA HOXA11 IKBKG MAML2 IFNAR2 CCNC CREM GPBP1 IRF7 ZBTB17 CRX RUVBL1 SOX9 FGFR1 EWSR1 SMARCB1 RPS6KA5 FOXO1 MAML3 KAT2A AP1B1 TP53 NFIC SS18L1 HOXB1 HLF H3F3A DDX5 CUX1 RUNX1 GABPA SRF CDC25B MSH6 SREBF2 TP73 MED6 CREBBP STAT4 NFATC4 ABCC9 NMI PROX1 HMGA1 ATF1 MED1 MAF GAK HOXA10 POU2F3 CHUK SERTAD2 CDK8 IKBKB RARA SMAD1 EIF2B1 SMARCA4 CITED4 NPAS2 KLF13 CNOT3 HDAC1 CITED2 MYBL2 MED21 GMEB1 TCF7L2 RXRG CTBP1 KLF4 UBTF NLK PIAS1 SERTAD3 NKX2-1 XRCC6 CEBPB KAT5 NAP1L1 HTT BRCA1 MAML1 EGR1 TAF7L RBPJ KHDRBS1 MDC1 ALX1 TGS1 RPA2 MED24 RPS6KA1 FOS PHOX2A PIAS3 MDM2 HNF1A ZNF639 ELK1 FOXM1 ETS2 MTDH CRTC2 NUP98 HOXA9 HIF1A ING1 PCMT1 AIRE SUV39H1 TCF3 PPARGC1A RPS6KA3 GMEB2 CDX2 ATF4 ETS1 KAT2B TRERF1 SND1 SH3GL1 PAX5 NCOA1 CREB1 TCF12 EP300 HSF1 DAXX TRIP10 HIPK2 AR PPARG NOTCH1 HOXD4 E2F3 POLR2A NOTCH3 MAST1 JUN CENPJ NFE2L2 RXRA N4BP2 DACH1 PPARA SMAD4 SRCAP CTNNB1 SMAD3 POU1F1 MYOD1 NEUROG1 NFATC2 PRKCD STAT2 H3F3B SNW1 MGMT MYBL1 MLL KLF1 STAT3 TRIM21 GATA1 HOXD10 MYC MECOM RBCK1 THRA CSNK2A2 DHX9 CDH2 NR3C1 E2F1 ATF3 PRRX2 E2F5 WRB NFE2 ESR1 CDKN1A KPNA2 MED15 YWHAH NOTCH2 GLI3 RPS6KA2 MYB HNF1B HNF4A CSK NFIA SPI1 AP2A2 MSH2 CARM1 TACC2 GTF2B HOXB2 RBBP7 MBD2 WT1 HOXB3 IRF3 HOXB6 FOXO4 ATF2 NOTCH4 HOXB9 STAT1 HMX3 STAT5A MSX1 GCM1 STAT5B GATA2 SMAD2 ABCA1 STAT6 CITED1 BCL3 MCM7 CDK5RAP3 CAMK4 PML SNIP1 TRIP4 TRAM2 ACTA2 MAFG*
Diabetes_McGill_ER_Stress	Diabetes_McGill_ER_Stress	*EIF2AK3 ERN1 ATF6 CCDC88B EIF2S1 XBP1 ATF4 ATF5 PPP1R15A DDIT3 BCL2 MAPK8 MBTPS1 MBTPS2 ATF6B NR0B2 DERL1 UCHL1 EIF2AK4 EIF2AK1 EIF2AK2 SYVN1 WFS1 HSPA5 ATP2A2 PDX1 MAFA SEC61A1 SEC61A2 SEC61B SEC61G SIAH1 SIAH2*
Diabetes_McGill_Inflamatory_Cytokines	Diabetes_McGill_Inflamatory_Cytokines	*AGT GPR77 CCR2 CCR5 CD74 CHUK CRADD CXCR1 CXCR2 CXCR3 CXCR4 ECSIT FADD FGF21 FGFR1 FGFR2 FGFR3 IKBKB IKBKE IKBKG IL10 IL10RA IL10RB IL13 IL13RA1 IL13RA2 IL1B IL1R1 IL6 IL6R IL6ST IL8 IRAK1 IRAK2 JAK1 JAK2 JAK3 KDR KLB LAP3 LTBP1 LTBP2 LTBP3 LTBP4 MADD MAP2K1 MAP2K2 MAP3K1 MAP3K14 MAP4K1 MAPK14 MAPK8 CCL2 NAMPT NFKB1 NFKBIA NMNAT1 NMNAT2 NMNAT3 PIK3C2A PIK3C2B PIK3CA RHOA RIPK1 RIPK2 RIPK3 RIPK4 SHC1 SHC2 SIRT1 SIRT2 SIRT3 SIRT4 SIRT5 SIRT6 SIRT7 SMAD2 SMAD3 SMAD4 SMAD6 SMAD7 STAT3 TAB1 MAP3K7 TGFB1 TGFB2 TGFB3 LEFTY2 TGFBR1 TGFBR2 TNF TNFRSF1A TNFRSF1B TOLLIP TRADD TRAF2 TRAF6 TYK2 VEGFA XIAP*
Diabetes_McGill_Insulin	Diabetes_McGill_Insulin	*STX1A STX1B SNAP25 VAMP2 SYT7 SYT5 SENP1 ATF6 XBP1 ERN1 EIF2AK3 ATF4 EIF2A*
Diabetes_McGill_Leptin_Pathway	Diabetes_McGill_Leptin_Pathway	*MAPK1 MAPK3 GRB2 IRS1 JAK1 JAK2 LEP LEPR PTPN11 SOCS3 STAT3 STAT5A*
Diabetes_McGill_Mtor_pathway	Diabetes_McGill_Mtor_pathway	*AKT1 EIF3A EIF4A1 EIF4A2 EIF4B EIF4E EIF4EBP1 EIF4G1 EIF4G2 EIF4G3 FKBP1A MKNK1 MTOR PDK2 PDPK1 PIK3CA PIK3R1 PPP2CA PTEN RPS6 RPS6KB1 TSC1 TSC2*
Diabetes_McGill_RAS	Diabetes_McGill_RAS	*DIRAS1 DIRAS2 DIRAS3 ERAS GEM HRAS KRAS MRAS NKIRAS1 NKIRAS2 NRAS RALA RALB RAP1A RAP1B RAP2A RAP2B RAP2C RASD1 RASD2 RASL10A RASL10B RASL11A RASL11B RASL12 REM1 REM2 RERG RERGL RRAD RRAS RRAS2 RASSF1 RASSF2 RASSF3 RASSF4 RASSF5 RASSF6 RASSF7 RASSF8 RASSF9 RASSF10*
Diabetes_McGill_Triggering_pathway_mediating_stimulation_of_insulin_secretion	Diabetes_McGill_Triggering_pathway_mediating_stimulation_of_insulin_secretion	*SLC2A1 SLC2A3 GCK PKLR PKM2 DLAT DLD PDHA1 PDHB PDHX PDP1 CS ACO1 ACO2 IDH2 OGDH DLST SUCLA2 SUCLG1 SUCLG2 SDHA SDHB SDHC SDHD FH MDH1 MDH2*
Diabetes_McGill_Wnt_pathway	Diabetes_McGill_Wnt_pathway	*APC AXIN1 BTRC CCND1 CREBBP CSNK1A1 CSNK1D CSNK2A1 CTBP1 CTNNB1 DVL1 FRAT1 FZD1 GSK3B HDAC1 LEF1 MAP3K7 MYC NLK PPARD PPP2CA SMAD4 TAB1 TLE1 WIF1 WNT1*
Mendelian_Blood_Desease	Mendelian_Blood_Desease	*ITGB2 FERMT3 SLC35C1 HBA1 HBA2 HBB RPL11 RPL35A RPL26 RPL5 RPS10 RPS17 RPS17L RPS19 RPS24 RPS26 RPS7*
Mendelian_Cerebral_Degeneration_Due_to_Generalized_Lipidoses	Mendelian_Cerebral_Degeneration_Due_to_Generalized_Lipidoses	*GLA NPC1 NPC2 SMPD1 GBA SLC37A4 GAA AGL PYGM PHKA1 PHKA2 PHKB PHKG2 LAMP2 PGAM2*
Mendelian_Disorders_of_Copper_Metabolism	Mendelian_Disorders_of_Copper_Metabolism	*ATP7A ATP7B PRNP*
Mendelian_Disorders_of_Fatty_Acid_Oxidation	Mendelian_Disorders_of_Fatty_Acid_Oxidation	*ETFA ETFB ETFDH CPT2 ACADSB ACADVL CPT1A HADHA SLC25A20 ACADM*

Mendelian_Etc	Mendelian_Etc	*HTT GFAP ARSA PSAP GALC SCN9A FXN MEFV SAA1 AR MEN1 RET*
Mendelian_Hereditary_Sensory_Neuropathy	Mendelian_Hereditary_Sensory_Neuropathy	*NTRK1 PMP22 MPZ LITAF EGR2 NEFL MFN2 KIF1B RAB7A LMNA TRPV4 BSCL2 GARS HSPB1 GDAP1 HSPB8 DNM2 MTMR2 SBF2 SH3TC2 NDRG1 PRX FGD4 FIG4 YARS GJB1 PRPS1 MED25 INF2 KARS*
Mendelian_Immune	Mendelian_Immune	*AIRE CD40LG FAS FASLG CYBB CYBA NCF1 NCF2 NCF4*
Mendelian_Long_QT_syndrome	Mendelian_Long_QT_syndrome	*KCNQ1 KCNH2 KCNE1 KCNE2 CACNA1C CAV3 SCN5A SCN4B*
Mendelian_Metabolism	Mendelian_Metabolism	*SLC22A5 APRT HPRT1 UMPS SLC25A15 NAGS CPS1 ASS1 ASL ARG1 CLCN5 DMP1 ENPP1 FGF23 PHEX SLC34A3 CYP27B1 ABCA1 APOA1 LPL APOB MTTP SAR1B LCAT GCDH PSPH DHTKD1 AHCY GNMT MAT1A GATM GLDC G6PD*
Mendelian_Severe_Combined_Immunodeficiency	Mendelian_Severe_Combined_Immunodeficiency	*IL2RG JAK3 ADA RAG1 RAG2 ZAP70 PNP NHEJ1 IL7R CD3D DCLRE1C PTPRC RFX5 RFXANK RFXAP AK2 CIITA*

Author contributions

Sample Collection And Phenotyping (WholeGenomes): L.L.B., J. Fadista, C. Herder, C.J.G., H.A.K., L.K., J. Kravic, V.L., C. Ladenvall, W.R., N.N., M.R., A. Swift, P.N., B.T., C. Meisinger, A.P., C. Huth, J. Tuomilehto, R.M.W., R.N.B., K.S.S., C.G., K.R.O., K. Strauch, T.T., A.T.H., F.S.C., T.D.S., T.M.F., N.P.B., K.L.M., L.G., M.B., D. Altshuler, M.I.M.

Sample Collection And Phenotyping (GWAS Imputation): C.F., H.C., M.M.-N., R.A.S., B.B., H.B., V.G., O.G., P.K., Y.L., C.N., D. Palli, D.R., D.T., Y.T.v.d.S., C. Ladenvall, E.M., A.-C.S., E.P.B., C. Langenberg, P. Froguel, R.J.F.L., A. Metspalu, N.J.W., E.I., T.I., P.W.F., R.M., L. Lind, J.B.M., L.J.S., M.B., D. Altshuler, M.I.M.

Sample Collection And Phenotyping (Whole Exomes): J.E.B., N.B., B.-G.H., C.P.J., T.K., J. Kuusisto, M.C.Y.N., N.D.P., A. Stančáková, H.E.A., U.A., D. Aguilar, R.A., T.A., E.C., C.-Y.C., A.C., J.E.C., V.S.F., S.P.F., B.I.F., M.G., D.E.H., P.J.H., C.-C.K., S.K., B.L., W.Y.L., J. Liu, M. Loh, S.K.M., S. Puppala, W.R.S., S.-T.T., H.A.T.Jr, F.T., G.W.Sr, T.Y.W., J.C.L., M.M., L.L.B., J. Fadista, G.L.S., C.J.G., L.K., D.H., J. Kravic, C. Ladenvall, N.N., A. Swift, P.N., S.P.O'R., J. Tuomilehto, Y.S.C., K.S.C., D.M.L., T.I.P., K.S.S., R.A.D., B.G., M.W., J.S.K., J.-Y.L., A.T.H., D.W.B., G. Atzmon, J.C.C., T.D.S., M. Laakso, G.I.B., J.B., R.D., E.S.T., C.L.H., J.G.W., T.M.F., N.J.C., L.G., M.B., D. Altshuler, M.I.M.

Sample Collection And Phenotyping (Exome Array & Replication): A. Mahajan, N.R.R., N.W.R., N.G., R.A.S., J.H., D. Pasko, T.V.V., S.-H.K., K.S.P., J.C.L., M.M., G.L.S., C.J.G., J.B.-J., I. Brandslund, C.C., A.S.F.D., T.E., A.J.F., L.H., D.H., J. Kravic, M. Hollensted, M.E.J., T.J., C. Ladenvall, J.M.J., A. Käräjämäki, L. Lannfelt, T.L., A.L., O.M., L. Milani, M.N., M.O.-M., L.Q., Q.Q., M.R., O.R., A.H.R., K. Stirrups, A.R.W., E.M., M.H.d.A., P. Deloukas, B.T., T.H., C. Meisinger, F.B.H., B.I., F.K., L. Liang, A.P., S.P.O'R., C.N.A.P., O.P., R.R., V.S., A.-C.S., R.N.B., C. Langenberg, K.S.S., A. Metspalu, N.J.W., M.W., C.G., E.I., T.I., P.W.F., R.M., L. Lind, K.R.O., K. Strauch, T.T., A.D.M., A.T.H., F.S.C., T.D.S., T.M.F., L.G., A.P.M., M.I.M.

Design And Study Supervision: C.F., T. Fingerlin, C. Hu, C.P.J., H.E.A., D. Prabhakaran, J.S., P.R.N., M.H.d.A., T.H., O.P., J. Tuomilehto, R.M.W., D. Bharadwaj, Y.S.C., G.R.C., J.C.N.C., K.S.C., M.J.D., S.B.E., P.E., K.A.J., D.M.L., W.J., R.C.W.M., T.I.P., M. Sandhu, N.T., P. Froguel, I. Barroso, Y.Y.T., E.Z., R.A.D., B.G., I.P., A.L.G., J.S.K., J.-Y.L., T.P., P. Donnelly, A.D.M., A.T.H., D.W.B., G. Atzmon, J.C.C., M. Laakso, T.M.S., G.I.B., J.B., R.D., E.S.T., G.M., C.L.H., J.G.W., M. Seielstad, T.M.F., J.B.M., N.J.C., R.S., E.S.L., N.P.B., K.L.M., T.M., L.G., G. Abecasis, J.C.F., L.J.S.,

A.P.M., M.B., D. Altshuler, M.I.M.

Data Production: J. Flannick, K.J.G., P. Fontanillas, M. Hollensted, J.M.J., C.B., J. Maguire, R.P., K. Shakir, T. Fennell, M.D., J. Murphy, R.O., J.S.R., H.G., E.B., G.B., J. Trakalo, D. Buck, Y.F., T.M.S., E.S.L., S.G., N.P.B., T.M.

Variant Calling And Panel Generation: C.F., J. Flannick, A. Mahajan, K.J.G., P. Fontanillas, M.A.R., X.S., N.R.R., N.W.R., P.S.C., C. Hartl, A.U.J., J.R.H., R.D.P., A. Kumar, M.M.-N., N.G., H.M.S., M.L.S., S.C.J.P., J.C., G.G., J.D.S., B.M.N., S. Purcell, T.S., T.W., J.B.-J., J. Kriebel, M.O.C., J. Maguire, R.P., K. Shakir, M.D., A.P.G., G.J., J. Murphy, R.O., J.S.R., E.B., Y.F., T.M.S., N.P.B., T.M., H.M.K., M.B., M.I.M.

Statistical Analysis: C.F., J. Flannick, T.M.T., A. Mahajan, V.A., K.J.G., C. Ma, P. Fontanillas, L. Moutsianas, D.J.M., M.A.R., J.R.B.P., X.S., T.W.B., N.R.R., N.W.R., P. Cingolani, A.E.L., J.F.T., H.M.H., J. Dupuis, P.S.C., C.M.L., C. Hartl, A.U.J., H.C., J.R.H., M.v.d.B., R.D.P., A. Kumar, M.M.-N., N.G., H.M.S., E.R.G., Jaehoon Lee, Y.C., R.A.S., J.E.B., P. Chen, J.H., M.J.G., D. Pasko, T.V.V., T.G., N.L.B., A.G.D.-W., T. Ferreira, M. Horikoshi, I.H., M.K.I., B.-J.K., Y.K., Y.J.K., M.-S.K., Juyoung Lee, S.L., K.-H.L., T.J.M., Y.N., X.W., R.P.W., J.Y., W. Zhang, N.B., B.F.V., B.-G.H., T.K., J. Kuusisto, A. Manning, M.C.Y.N., N.D.P., B.B., A. Stančáková, A.S.B., J.M.M.H., H.M.L., Y.L., W. Zhao, J. Danesh, V.K.L.L., D.S., W.Y.S., C.H.T.T., L.Y., C. Ladenvall, J.C.N.C., C. Langenberg, R.C.W.M., T.I.P., R.J.F.L., J.S.R., N.J.W., E.I., H.K.I., P.W.F., I.P., R.M., K. Strauch, J.B.M., R.S., N.P.B., K.L.M., G. Abecasis, L.J.S., A.P.M., H.M.K., M.B., D. Altshuler, M.I.M.

Drafting Of Manuscript: C.F., J. Flannick, T.M.T., A. Mahajan, V.A., K.J.G., P. Fontanillas, L. Moutsianas, D.J.M., M.A.R., J.R.B.P., P. Cingolani, N.J.C., R.S., N.P.B., L.J.S., A.P.M., H.M.K., M.B., D. Altshuler, M.I.M.

Project Leadership: M.B., D. Altshuler, M.I.M.

Acknowledgements (By Lead Institution)

Albert Einstein College of Medicine, USA

Funders of this work:

- The American Federation for Aging Research
- The Einstein Glenn Center
- National Institute on Aging (PO1AG027734, R01AG046949, 1R01AG042188, P30AG038072)

Broad Institute, USA

Funders of this work:

- NHGRI ("Large Scale Sequencing and Analysis of Genomes" U54HG003067)
- NIDDK ("Multiethnic Study of Type 2 Diabetes Genes" U01DK085526)
- NIH ("Low-Pass Sequencing and High Density SNP Genotyping in Type 2 Diabetes" 1RC2DK088389)

Personal support:

- Vineeta Agarwala: National Institute of General Medical Sciences award (T32GM007753)

Center for Genome Science, National Institute of Health, Republic of Korea

Funders of this work:

- Korea National Institute of Health (2012-N73002-00)
- Korea National Institute of Health and Korea Centers for Disease Control and Prevention (4845–301)

Other acknowledgements:

- This study was provided with biospecimens and data from the Korean Genome Analysis Project (4845-301), the Korean Genome and Epidemiology Study (4851-302), and the Korea Biobank Project (4851-307, KBP-2013-11 and KBP-2014-68) that were supported by the Korea Centers for Disease Control and Prevention, Republic of Korea.

The Chinese University of Hong Kong

Funders of this work:

- The Focused Investment Scheme of the Chinese University of Hong Kong
- The Hong Kong Foundation for Research and Development in Diabetes established under the auspices of the Chinese University of Hong Kong
- Hong Kong Governments Research Grant Committee Central Allocation Scheme (CUHK 1/04C)
- The Innovation and Technology Fund (ITS/487/09FP, ITS/130/11)
- The Honk Kong Research Grants Council Theme-based Research Scheme (T12-402/13N)

German Diabetes Center, Germany

Funders of this work:

- Ministry of Science and Research of the State of North Rhine-Westphalia (MIWF NRW)
- German Federal Ministry of Health (BMG)
- Grant from the German Federal Ministry of Education and Research (BMBF)

GoT2D consortium

Funders of this work:

- National Institutes of Health ("Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes" RC2DK088389)
- The German Center for Diabetes Research (DZD)

Hallym University Chuncheon, Republic of Korea

Funders of this work:

- National Research Foundation of Korea (NRF-2012R1A2A1A03006155)

Helmholtz Zentrum München – German Research Center for Environmental Health, Germany

Funders of this work:

- The German Center for Diabetes Research (DZD)
- Helmholtz Zentrum München (German Research Center for Environmental Health), which is supported by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria
- The Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ

Other acknowledgements:

- The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

Imperial College London, UK

Funders of this work:

- Action on Hearing Loss (G51)
- The British Heart Foundation (SP/04/002)
- European Union FP7 (EpiMigrant, 279143)
- Medical Research Council (G0601966, G0700931)
- MRC-PHE Centre for Environment and Health
- The National Institute for Health Research (NIHR) (RP-PG-0407-10371)
- NIHR Biomedical Research Centre at Imperial College Health Care NHS Trust
- NIHR Health Protection Research Unit on Health Impact of Environmental Hazards
- The Wellcome Trust (084723)

Personal Support

- Paul Elliot: NIHR Senior Investigator

Other acknowledgements:

- The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust. The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. We thank the participants and research staff who made the study possible.

The Jackson Heart Study, USA

Funders of this work:

- National Heart, Lung, and Blood Institute and the National Institute on Minority Health a nd Health Disparities (HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C)

The Jackson Laboratory for Genomic Medicine, USA

Funders of this work:

- National Institutes of Health (R00DK092251)

King's College, London

Funders of this work:

- European Community's Seventh Framework Programme (FP7/2007-2013)
- National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London
- The Wellcome Trust

Lund University, Sweden

Funders of this work:

- The Academy of Finland (263401, 267882)
- European Research Council Advanced Research Grant
- The Folkhälsan Research Foundation
- Novo Nordisk
- The Påhlssons Foundation
- The Sigrid Juselius Foundation
- The Skåne Regional Health Authority
- The Swedish Heart-Lung Foundation
- Swedish Research Council (Linné and Strategic Research Grant)

Massachusetts General Hospital, USA

Funders of this work:

- National Institutes of Health (U01DK085526)

Personal support:

- Jose Florez: MGH Research Scholar
- James B Meigs: National Institutes of Health (K24DK080140)

McGill University, Canada

Funders of this work:

- The Canadian Institutes of Health Research

Personal support:

- Rob Sladek: Chercheur Boursier award from the Fonds de la Recherche en Santé du Québec; New Investigator Award from the Canadian Institutes of Health Research

National Institute for Health and Welfare, Helsinki, Finland

Funders of this work:

- The Academy of Finland (139635)
- The Finnish Foundation for Cardiovascular Research

National University of Singapore

Funders of this work:

- Biomedical Research Council (BMRC) Individual Research Grant
- National Medical Research Council (NMRC) Individual Research Grant
- NMRC Centre Grant

Personal support:

- Ching-Yu Cheng: NMRC Clinician Scientist award
- E Shyong Tai: NMRC Clinician Scientist award
- YY Teo: National Research Foundation Fellowship
- TY Wong: NMRC Singapore Translational Research Investigator award

Seoul National University, Republic of Korea

Funders of this work:

- Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI14C0060)
- National Research Foundation of Korea (NRF) grants (2012R1A3A2026438, 2013M3A9C4078158, 2008-0062618)

T2D-GENES and GoT2D consortium

Funders of this work:

- National Institutes of Health ("Multiethnic Study of Type 2 Diabetes Genes" U01s DK085526, DK085501, DK085524, DK085545, DK085584; "Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes" DK088389)
- The German Center for Diabetes Research (DZD)

University of Bergen, Norway

Funders of this work:

- European Research Council (ERC-2011-ADG_20110310 #293574)
- The KG Jebsen Foundation
- Research Council of Norway
- University of Bergen
- The Western Norway Regional Health Authority (Helse Vest)

University of Chicago, USA

Funders of this work:

- National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (P30DK020595)
- National Institutes of Health (R01MH101820, P60DK20595, U01DK085501, R01HL102830, U01HG005773, R01MH090937)

Personal support:

- Hae Kyung Im: National Cancer Institute (K12CA139160) awarded by the Institute of Translational Medicine at the University of Chicago

Other acknowledgements:

- The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

University of Cambridge, UK

Funders of this work:
- UK Medical Research Council (G0800270)
- British Heart Foundation (SP/09/002)
- European Research Council (268834)
- European Union Framework Programme 7 (HEALTH-F2-2012-279233)
- National Institute for Health Research, Cambridge Biomedical Research Centre
- Merck
- Pfizer

Personal support:
- John Danesh: British Heart Foundation Professorship, NIHR Senior Investigator, European Research Council Senior Investigator.

University of Copenhagen

Funders of this work:
- The Lundbeck Foundation grant to LuCamp (www.lucamp.org)
- The Novo Nordisk Foundation Center for Basic Metabolic Research, an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation

University of Eastern Finland

Funders of this work:
- The Academy of Finland
- Strategic Research Funding from the University of Eastern Finland, Kuopio, Finland

University of Helsinki, Finland

Personal support:
Heikki A Koistinen: Academy of Finland Clinical Researcher

University of Michigan, USA

Funders of this work:
- National Institutes of Health (R01DK062370, R01DK098032, RC2DK088389)

University of North Carolina, USA

Funders of this work:
- National Institutes of Health (DK072193, DK093757)

University of Oxford, UK

Funders of this work:
- The European Commission (ENGAGE: HEALTH-F4-2007-201413; Marie-Curie Fellowship PIEF-GA-2012-329156)
- MRC (G0601261, G0900747-91070)
- National Institutes of Health (RC2-DK088389, DK085545, DK098032)
- Wellcome Trust (064890, 083948, 085475, 086596, 090367, 090532, 092447, 095101, 095552, 098017, 098381, 100956)
- Diabetes UK (for the BDA Warren repository)

Personal support:
- Peter Donnelly: Wellcome Trust Senior Investigator

- Andrew Farmer: NIHR Senior Investigator
- Anna L Gloyn: Wellcome Trust Senior Fellow in Basic Biomedical Science
- Fredrik Karpe: NIHR Oxford Biomedical Research Centre; NIHR National Bioresource
- Cecilia Lindgren: Wellcome Trust Intermediate Research Fellow; Li Ka Shing Foundation
- Davis J McCarthy: General Sir John Monash Scholarship
- Mark I McCarthy: Wellcome Trust Senior Investigator; NIHR Senior Investigator
- Gilean McVean: Wellcome Trust Senior Investigator
- Andrew P Morris: Wellcome Trust Senior Fellow in Basic Biomedical Science
- Katharine R Owen: NIHR Clinician Scientist
- John R B Perry: Sir Henry Wellcome Postdoctoral Fellowship
- Manuel A Rivas: NDM Prize Studentship - Clarendon Award
- Juan Fernandez Tajes: Marie-Curie Fellow
- Martijn van de Bunt: NDM Prize Studentship

Other acknowledgements:

- We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for the generation of array and sequencing data.

The University of Texas Health Science Center at Houston, USA

Funders of this work:

- National Institutes of Health (U01DK085501, R01HL102830, R01DK073541)

Uppsala University

Funders of this work:

- The European Research Council
- The Knut och Alice Wallenberg Foundation
- The Swedish Heart-Lung Foundation (20140422)
- The Swedish Diabetes Foundation (2013-024)
- The Swedish Research Council (2012-1397)
- Uppsala University

Wake Forest School of Medicine, USA

Funders of this work:

- National Institutes of Health (R01DK066358)

Wellcome Trust Sanger Institute, UK

Funders of this work:

- National Institute for Health Research
- The Wellcome Trust (098051)

Other acknowledgements:

- Panos Deloukas's work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit, which is supported and funded by the National Institute for Health Research.

Studies

EGCUT received financing from European Regional Development Fund, road-map grant no.3.2.0304.11-0312 and grant "Center of Excellence in Genomics (EXCEGEN). EGCUT studies were covered also by targeted financing from Estonian Government (IUT24-6, IUT20-60) and CTG grant (SP1GVARENG) from Development Fund of the University of Tartu.

The Botnia study has been financially supported by grants from the Sigrid Juselius Foundation, Folkhälsan Research Foundation, Nordic Center of Excellence in Disease Genetics, an EU grant (EXGENESIS), Signe and Ane Gyllenberg Foundation, Swedish Cultural Foundation in Finland, Finnish Diabetes Research Foundation, Foundation for Life and Health in Finland, Finnish Medical Society, Paavo Nurmi Foundation, Helsinki University Central Hospital Research Foundation, Perklén Foundation, Ollqvist Foundation, Närpes Health Care Foundation and Ahokas Foundation. The study has also been supported by the Ministry of Education in Finland, Municipal Heath Care Center and Hospital in Jakobstad and Health Care Centers in Vasa, Närpes and Korsholm.

The Mount Sinai BioMe Biobank has been financially supported The Andrea and Charles Bronfman Philanthropies.

Other
This study also utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov).

Current addresses

- David Altshuler: Vertex Pharmaceuticals, Boston, MA, USA
- Rector Arya, Ravindranath Duggirala, Vidya S Farook, Christopher P Jenkinson: South Texas Diabetes and Obesity Institute, Edinburg Regional Academic Health Center, University of Texas Rio Grande Valley, Edinburg, TX 78541, USA
- Dwaipayan Bharadwaj: School of Biotechnology, Jawaharlal Nehru University, New Delhi
- Giriraj Chandak: Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, Telangana, India
- Nancy J Cox: Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA
- Aaron Day-Williams: Biogen, Boston, MA, USA
- Mark DePristo: Verily Life Sciences, Mountain View, CA 94043, USA
- Pierre Fontanillas: 23andMe, Mountain View, CA 94041, USA
- Eric R Gamazon: Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; Department of Clinical Epidemiology and Biostatistics and Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
- Heather Highland: Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA
- Jinyan Huang: State Key Laboratory of Medical Genomics, Shanghai Institute of Hematology, Rui Jin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
- Jeroen Hyghe: Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA
- Erik Ingelsson: Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA
- Claes Ladenvall: Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden
- Cecilia M Lindgren: The Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford OX3 7BN, UK
- Jong-Young Lee: NGS Division/BIO R&D, 2nd Floor, B-dong, AICT Bldg., 145 Gwanggyo-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do,443-270 Republic of Korea
- Adam Locke: McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO, USA
- Marie Loh: Translational Laboratory in Genetic Medicine (TLGM), National University of Singapore (NUS) and A*STAR, Singapore
- Clement Ma: Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Boston, MA, USA
- Taylor J. Maxwell: Computational Biology Institute, The George Washington University, Ashburn, Virginia, USA.
- Loukas Moutsianas: Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK
- Lu Qi: Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA
- Xueling Sim: Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore
- Kathleen Stirrups: Department of Haematology, University of Cambridge, Cambridge, UK
- Farook Thameem: Department of Biochemistry, Faculty of Medicine, Kuwait University, Kuwait
- Jaakko Tuomilehto: Dasman Diabetes Institute, Dasman, Kuwait