Novel DNA methylation sites of glucose and insulin homeostasis: an integrative cross-omics analysis

Jun Liu[1*], Elena Carnero-Montoro[1,2,3*], Jenny van Dongen[4], Samantha Lent[5], Ivana Nedeljkovic[1], Symen Ligthart[1], Pei-Chien Tsai[3,6,7], Tiphaine C. Martin[3,8,9], Pooja R. Mandaviya[10], Rick Jansen[11], Marjolein J. Peters[10], Liesbeth Duijts[12,13], Vincent W.V. Jaddoe[1,14,15], Henning Tiemeier[1,14,16], Janine F. Felix[1,14,17], Audrey Y Chu[18,19], Daniel Levy[18,19], Shih-Jen Hwang[18,19], Jan Bressler[20], Rahul Gondalia[21], Elias L. Salfati[22], Christian Herder[23,24], Bertha A. Hidalgo[25], Toshiko Tanaka[26], Ann Zenobia Moore[26], Rozenn N. Lemaitre[27], Min A. Jhun[28], Jennifer A. Smith[28], Nona Sotoodehnia[27], Stefania Bandinelli[29], Luigi Ferrucci[26], Donna K. Arnett[30], Harald Grallert[23,31], Themistocles L. Assimes[22], Lifang Hou[32,33], Andrea Baccarelli[34], Eric A Whitsel[35,36], Ko Willems van Dijk[37,38], Najaf Amin[1], André G. Uitterlinden[1,10], Eric J.G. Sijbrands[10], Oscar H. Franco[1,39], Abbas Dehghan[1,40], Tim D. Spector[3], Josée Dupuis[5], Marie-France Hivert[41,42,43], Jerome I. Rotter[44], James B. Meigs[45,46,47], James S. Pankow[48], Joyce B.J. van Meurs[10], Aaron Isaacs[1,49], Dorret I. Boomsma[4], Jordana T. Bell[3], Ayşe Demirkan[1,37§] and Cornelia M. van Duijn[1,50,51§]

[*] These authors contributed equally to this work.

[§] These authors jointly supervised this work.

Affiliations:

1 Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

2 Center for Genomics and Oncological Research, GENYO, Pfizer/University of Granada/Andalusian Government, PTS, Granada, Spain

3 Department of Twin Research and Genetic Epidemiology, King's College London, London, UK

4 Department of Biological Psychology, Amsterdam Public Health (APH) research institute, VU University Amsterdam, Amsterdam, the Netherlands

5 Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

6 Department of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan

7 Division of Allergy, Asthma, and Rheumatology, Department of Pediatrics, Chang Gung Memorial Hospital, Linkou, Taiwan

8 Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

9 The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

10 Department of Internal Medicine, Genetics Laboratory, Erasmus University Medical Center, Rotterdam, the Netherlands

11 Department of Psychiatry, VU University Medical Center, Amsterdam, the Netherlands

12 Division of Neonatology, Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands

13 Division of Respiratory Medicine, Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands

14 Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands

15 The Generation R Study Group, Erasmus University Medical Center, Rotterdam, the Netherlands

16 Department of Child and Adolescent Psychiatry, Erasmus University Medical Center, Rotterdam, the Netherlands

17 Generation R Study Group, Erasmus University Medical Center, Rotterdam, the Netherlands

18 The Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

19 The Framingham Heart Study, National Heart, Lung and Blood Institute, National Institutes of Health, Framingham, MA, USA

20 Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA.

21 Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

22 Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

23 German Center for Diabetes Research (DZD), München- Neuherberg, Germany

24 Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany

25 Department of Epidemiology, University of Alabama at Birmingham, AL, USA

26 Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA

27 Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA

28 Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

29 Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy

30 School of Public Health, University of Kentucky, Lexington. KY, USA

31 Research Unit of Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München Research Center for Environmental Health, Neuherberg, Germany

32 Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University Chicago, Evanston, IL, USA

33 Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

34 Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA

35 Department of Epidemiology, University of North Carolina Gillings School of Global Public Health, Chapel Hill, USA

36 Department of Medicine, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA

37 Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands

38 Department of Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands

39 Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

40 Department of Epidemiology and Biostatistics, Imperial College London, London, UK

41 Department of Medicine, Université de Sherbrooke, Sherbrooke, QC, Canada

42 Diabetes Unit, Massachusetts General Hospital, Boston, MA, USA

43 Division of Chronic Disease Research Across the Lifecourse, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

44 The Institute for Translational Genomics and Population Sciences and Departments of Pediatrics and Medicine, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA

45 Department of Medicine, Harvard Medical School, Boston, MA, USA

46 Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA, USA

47 Programs in Metabolism and Medical & Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

48 Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

49 CARIM School for Cardiovascular Diseases, Maastricht Centre for Systems Biology (MaCSBio), and Dept. of Biochemistry, Maastricht University, Maastricht, the Netherlands

50 Leiden Academic Center for Drug Research, Leiden University, Leiden, the Netherlands

51 Nuffield Department of Population Health, Oxford University, Oxford, UK

Keywords:

Glucose, insulin, type 2 diabetes, omics, epigenomics, transcriptomics

Corresponding author: Cornelia M. van Duijn

Email address: cornelia.vanduijn@ndph.ox.ac.uk

Abstract

Despite existing reports on differential DNA methylation in type 2 diabetes (T2D) and obesity, our understanding of the functional relevance of the phenomenon remains limited. Because obesity is the main risk factor for T2D and a driver of methylation from previous study, we aimed to explore the effect of DNA methylation in the early phases of T2D pathology while accounting for body mass index (BMI). We performed a blood-based epigenome-wide association study (EWAS) of fasting glucose and insulin among 4,808 non-diabetic European individuals and replicated the findings in an independent sample consisting of 11,750 non-diabetic subjects. We integrated blood-based *in silico* cross-omics databases comprising genomics, epigenomics and transcriptomics collected by BIOS project of the Biobanking and BioMolecular resources Research Infrastructure of the Netherlands (BBMRI-NL), the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC), the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) consortium, and the tissue-specific Genotype-Tissue Expression (GTEx) project. We identified and replicated nine novel differentially methylated sites in whole blood (P-value < 1.27×10^{-7}): sites in *LETM1, RBM20, IRS2, MAN2A2* genes and 1q25.3 region were associated with fasting insulin; sites in *FCRL6*, *SLAMF1, APOBEC3H* genes and 15q26.1 region were associated with fasting glucose. The association between *SLAMF1*, *APOBEC3H* and 15q26.1 methylation sites and glucose emerged only when accounted for BMI. Follow-up *in silico* cross-omics analyses indicate that the *cis*-acting meQTLs near *SLAMF1* and *SLAMF1* expression are involved in glucose level regulation. Moreover, our data suggest that differential methylation in *FCRL6* may affect glucose level and the risk of T2D by regulating *FCLR6* expression in the liver. In conclusion, the present study provided nine new DNA methylation sites associated with glycemia homeostasis and also provided new insights of glycemia related loci into the genetics, epigenetics and transcriptomics pathways based on the integration of cross-omics data *in silico*.

Background

Type 2 diabetes (T2D) is a common metabolic disease, characterized by disturbances in glucose and insulin metabolism, that are in part genetically driven[1-10] with the heritability ranging from 20% to 80%[11]. DNA methylation has been associated with T2D as well as with fasting glucose and insulin[12]. Methylation-based risk scores of T2D predicted incident T2D cases that go beyond traditional risk factors such as obesity and waist-hip ratio[13]. Further, obesity, which is the most important determinant of insulin resistance and glucose levels in the population,[14,15] has also been associated with differential DNA methylation[13]. This raises the possibility that differential methylation associated with glucose and insulin levels could be counfounded by obesity. DNA methylation, mainly depending on the region, results in gene silencing and thus regulates gene expression and subsequent cellular functions[16]. It is very well possible that the epigenetic modifications occur in early phases of the pathology of T2D, requiring research focusing on the early process of the disease, e.g. in subjects free of diabetes.

We aimed to determine the association of DNA methylation with fasting glucose and insulin accounting for the effect of obesity in the non-diabetic subjects and to evaluate the impact of DNA methylation on cross-omics level. We followed the hypothesis that genetic variants drive DNA methylation which subsequently regulates gene expression and then glycemic traits, changes of which mark the early phases of diabetes pathology (**Figure 1a**). First, we performed a blood-based epigenome-wide association study (EWAS) meta-analysis of 4,808 diabetes-free individuals of European descent and replicated our findings among 11 cohorts summing up to 11,750 trans-ethnic non-diabetic individuals, mainly from European ancestry. Subsequently, we explored the role of genetics in determining the regulation of methylation associated with glycemic traits and the effects of the differential methylation on the human transcriptome *in silico* (**Figure 1a**).

Results

1. Blood-based epigenome-wide association analysis of glycemic traits

The discovery phase was based on four European cohorts (**Supplementary Table 1).** The meta-analysis revealed DNA methylation in 28 unique CpG sites associated with fasting glucose (11 CpG sites, n = 4,808) and/or insulin (20 CpG sites, n = 4,740) at epigenome-wide significance (P-value < 1.27×10^{-7}) in either the baseline model without body mass index (BMI) adjustment or in the second model with BMI adjustment. Of these 28 CpG sites, 15 were novel (**Table 1**) while 13 were identified by earlier EWAS studies of either T2D or related traits, including glucose, insulin, hemoglobin A1c (HbA1c), homeostatic model assessment-insulin

resistance (HOMA-IR) and BMI[12,13,17-25] (**Supplementary Table 2**). Of the known CpG sites, three located in *SLC7A11, CPT1A,* and *SREBF1* associated with both glucose and insulin. The remaining ten CpG sites associated with insulin and were located in genes *ASAM, DHCR24, RNF145, KDM2B, MYO5C, TMEM49, CPT1A, two in ABCG1* and one in the 4p15.33 region.

The 15 novel CpGs were tested using the same models in meta-analysis of 11 independent cohorts including 11,750 non-diabetic subjects from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (**Supplementary Table 1**). As a result, nine unique CpG-trait associations were replicated: four passing the epigenome-wide significance threshold (P-value < 1.27×10^{-7}) and five passing the Bonferroni significance threshold after correcting for 15 tests (P-value < 3.3×10^{-3}) (**Table 1**). These included five sites associated with fasting insulin in the baseline model (*LETM1*, *RBM20*, *IRS2, MAN2A2* and 1q25.3 region), one associated with fasting glucose in the baseline model (*FCRL6)* and additional three emerged to be associated with fasting glucose in the BMI-adjusted model (*SLAMF1, APOBEC3H* and 15q26.1 region).

Because the replication cohorts also included other ethnic groups than the main European ancestry (European: n = 7,254, African: n = 3,744, and Hispanic: n = 543), we also performed meta-analysis stratified by ancestry. Seven out of nine new CpG sites (*FCRL6, LETM1, RBM20, IRS2, MAN2A2, APOBEC3H* and the 15q26.1 region) confirmed consistent directions of effect across the three ethnicities. (**Supplementary Table 3)**

2. Integrated *in silico* cross-omics studies

To evaluate the functional relevance of differential methylation findings, we integrated our EWAS findings with genomics, epigenomics and transcriptomics data obtained from public resources. These included blood-based *cis* and *trans* methylation quantitative trait loci (meQTLs), expression quantitative trait methylations (eQTMs), expression quantitative trait loci (eQTLs) from the European BIOS database[26] from the Biobanking and BioMolecular resources Research Infrastructure of the Netherlands (BBMRI-NL), the genome-wide association studies (GWAS) of glycemic traits or T2D from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) and the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) consortium[4-7,27], and tissue-specific eQTL-phenotype associations from MetaXcan database[28,29] based on Genotype-Tissue Expression (GTEx) project (See resources of these database in **URLs**). The hypotheses tested were outlined in **Figure 1a**: DNA methylation and gene expression are partly genetically driven and heritable; genetic variants determine in part methylation, which subsequently influences expression and further fasting glucose and/or insulin. While doing so, we centered on the 11 top independent DNA methylation sites previously identified

(cg00574958 in *CPT1A* and cg06500161 in *ABCG1* were used) and 9 novel sites from our current study (total n = 20 methylation sites).

2.1 Genomics of the differentially methylated sites involved in glycemic traits

Using BIOS database (blood-based)[26], we found that 2,991 single-nucleotide polymorphisms (SNPs) in 29 unique genetic loci were associated with methylation in either *cis* or *trans* across 18 unique CpG sites among the tested 20 target methylation sites. For two CpG sites located in *SLC7A11* and *LETM1*, we did not find any significant meQTLs. Results are shown in **Figure 2** and given in detail in **Supplementary Table 4**. Seven of the 29 meQTLs, 5 *cis-acting* and 2 *trans-acting* were found significantly associated with T2D, fasting glucose or HbA1c (shown in **Figure 2** and given in detail **Supplementary Table 5**).

Based on our leading hypothesis, we examined whether DNA methylation may influence fasting glucose and insulin in the circulation. To this end, we performed a two-sample-based Mendelian Randomization (MR) analyses[30] to examine the causal effect of the differential DNA methylation sites in blood on fasting glucose or insulin using the summary GWAS results from BIOS[26] and MAGIC databases[5] (**Supplementary Table 6**). Up to eight independent genetic variants were included in the genetic risk score as the instrumental variable of each methylation site to check the association with the observationally associated traits, either fasting glucose or fasting insulin. Thirteen CpG sites out of the initial 20 met the present MR criteria and were testes by MR. No significant associations were detected when adjusting for multiple testing involving 13 independent tests (P-value < 3.8×10^{-3}) except for two marginal siginificant associations between methylation site in *RBM20* with fasting insulin (P-value = 0.04) and methylation site in *SLAMF1* with fasting glucose (P-value = 0.05).

2.2 Transcriptome associated with the differentially methylated sites of glycemic traits

2.2.1 Association of gene expression with differentially methylated sites in blood

To understand if the methylation is also eQTM, we investigated the association between gene expression and the 20 key glycemic methylation sites from the European blood-based BIOS database[26] (intergrated in **Figure 2**). We found that methylation in five CpG sites, including two novel sites (in *FCRL6* and *SLAMF1*) and three known sites (in *CPT1A*, *SREBF1* and *ABCG1*), was significantly negatively associated with the expression of their respective genes: *FCRL6* (P-value = 4.0×10^{-39}), *SLAMF1 (*P-value = 4.1×10^{-5}), *CPT1A* (P-value = 3.1×10^{-20}), *SREBF1* (P-value = 4.5×10^{-15}) and *ABCG1* (P-value = 2.2×10^{-37}). The methylation site in *SLAMF1* was positively associated with expression of two other genes near *SLAMF1* (*CD244:* P-value = 2.9×10^{-6} and *SLAMF7:* P-value = 5.4×10^{-9}). (**Supplementary Table 7**)

2.2.2 Common genetic determinants of glycemia related to DNA methylation and gene expression in blood

We next explored if the genetic variants associated with the differential expression above were the same as the meQTLs using the eQTL data from the European blood-based BIOS database[26] (integrated in **Figure 2** and detailed in **Supplementary Table 8**). We found three genetic determinants associated with both differential DNA methylation, including two novel methylation sites (in *FCRL6* and *SLAMF1*) and one known (in *SREBF1*), and gene expression in blood. Rs1577544 near *SLAMF1* associated with decreased methylation of the cg18881723 (Z = -5.45, P-value = 5.1×10^{-8}) as well as *SLAMF1* expession (Z = -6.40, P-value = 1.6×10^{-10}). Rs11265282 in *FCRL6* associated with increased methylation of cg00936728 (Z = 4.17, P-value = 3.0×10^{-5}) but decreased *FCRL6* expession (Z = -6.73, P-value = 1.7×10^{-11}). Rs6502629 in *TOM1L2* associated with increased methylation of cg11024682 (Z = 9.97, P-value = 2.1×10^{-23}) but decreased *SREBF1* expession (Z = -17.93, P-value = 7.2×10^{-72}).

2.2.3 Tissue-specific differential expression associated with T2D and related traits

We then explored the tissue-specific differential expression associated with T2D and related traits by data-mining from MetaXcan database from the GTEx project[28,29]. This analyses targeted on the eQTM-related expression of seven genes as shown in 2.2.1 in six glucose-metabolism-related tissues including blood, adipose subcutaneous, adipose visceral omentum, liver, pancreas, and muscle skeletal (**Supplementary Table 9**). The effect direction consistency was checked between methylation sites, gene expression and T2D or related traits. That meant the direction of the association between methylation and T2D or related traits should be a combination of the directions of methylation with gene expression and gene expression with T2D or related traits. The expression of *SREBF1* in blood was significantly associated with decreased levels of HbA1c (Z = -3.26, P-value = 1.1×10^{-3}) and also with decreased risk of T2D (Z = -2.40, P-value = 0.016). Meanwhile, the known cg11024682 in *SREBF1* was positively associated with fasting glucose (Z = 6.45, P-value = 2.7×10^{-8}) and fasting insulin (Z = 6.27, P-value = 6.7×10^{-9}) and negatively associated with expression of *SREBF1* in blood (Z = -7.84, 4.5×10^{-15}) (shown in **Figure 1c**). Higher liver gene expression of *FCRL6*, a novel locus, was associated with increased risk of T2D (Z = 2.14, P-value = 0.032) based on MetaXcan[28,29,31] results generated by integrating functional data in liver[32,33] and the GWAS of T2D[9]. The novel cg00936728 in *FCRL6* was negatively associated with fasting glucose (Z = -6.17, P-value = 9.1×10^{-8}) and expression of *FCRL6* in blood (Z = -13.09, 4.0×10^{-39}) (shown in **Figure 1b**).

Discussion

The current large-scale EWAS identified and replicated nine new methylation sites associated with fasting glucose or insulin, including three additionally uncovered sites (in *SLAMF1*, *APOBEC3H* and the 15q26.1 region) associated with fasting glucose only after adjustment for BMI. We further validated 13 previous reported CpG sites in 11 independent loci. Based on the cross-omics analyses, our report complements earlier studies[12,13,17-25,34] for multiple DNA methylation sites related to the pathology early in the development of T2D through genetics and/or gene expression. We also present *in silico* evidence supporting the potential involvement of the nine new methylation sites.

The novel methylation sites annotated to genes that play roles in glucose and energy metabolism (*IRS2*), metabolism of proteins (*MAN2A2* and *EDEM3*, the nearest gene of cg13222915), RNA and splicing regulation (*RBM20*), RNA metabolism (*APOBEC3H*), small molecule transport (*LETM1*) and immune system process (*SLAMF1, FCRL6* and *SV2B*, the nearest gene of cg18247172). Some of these genes are also involved in other diseases or biomarkers, including inflammatory phenotypes (*EDEM3* with systemic lupus erythematosus[35], *SLAMF1* with inflammatory bowel disease[36] and *FCRL6* with C-reactive protein (CRP)[37,38]), cardiovascular phenotypes (*RBM20* with electrocardiographic traits[39]), cancer (*IRS2* with prostate cancer[40]) and schizophrenia (*MAN2A2*)[41]. Thus, observations provided insight into the pathways that might link T2D to inflammation, cardiovascular disease, cancer and schizophrenia, all disorders associated epidemiologically or clinically with T2D. This phenomenon may point at genetic pleiotropy of the genes, i.e. a gene codes the same products in various cells or have cascade-like signaling function that affects various targets.

In this paper, we used the assumption that genetic variants drive DNA methylation which subsequently regulates gene expression and then glycemic traits[42]. Two pathways (on *SREBF1* and *FCRL6*) related to genetics-epigenetics-transcriptomics-phenotype were observed in the present study (**Figure 1b** and **Figure1c).** We validated the differential methylation of *SREBF1* in insulin metabolism[43] and extended the findings building a pathway based on the cascading cross-omics analysis in the assumption of genetics-epigenetics-transcriptomics-phenotype. We also discovered a new pathway of *FCRL6* in glucose metabolism, which still needs further research for its role to be fully understood. From the present study, with the integration of all the significant associations, the effect allele (*C*) of the genetic variant rs11265282 in *FCRL6* increases the methylation level which associates with lower expression of *FCRL6* in the blood. The decreased *FCRL6* expression in liver was also associates with decreased risk of T2D. This presumably is mediated by a decrease in fasting glucose.

The present study provides new genomic targets for further work on the pathology of T2D through large-scale EWAS and replication. However, the main findings are based on data from blood which was the only accessible tissue and may not be representative of more glucose-relevant tissues, although concordance of differential methylation between blood and adipose is high for certain pathways[44]. Our present MR analyses yielded no evidence for causality between methylation sites and fasting glucose or insulin. One limitation we faced here was the limited data to perform MR in all the association steps, e.g. the association of methylation with gene expression, the gene expression with phenotypes, some CpG sites with phenotypes, as well as the inverse causal effect of glucose or insulin on DNA methylation, thus we can not exclude entirely the influence of glycemia homeostasis on methylation levels. On the other hand, some of the MR tests performed had low explained variance of the instrumental variables, i.e. seven of the 13 performed CpGs have instrumental variables explained variance less than 5%. This might partly explain the insignificant findings in MR in the current study. Further studies are needed to include additional biologically relevant tissues and perform MR based on the tissue specific meQLTs.

In conclusion, our large-scale EWAS and replication have identified nine new differentially methylated sites associated with fasting glucose or insulin. The integrative *in silico* cross-omics analysis provided new insights of both known and new glycemia related loci into the genetics, epigenetics and transcriptomics pathway. Our study suggests that the expression of seven genes associated with either glycemia related DNA methylation is altered. Two of these seven expressed genes are also associated with T2D or related traits through the tissue-specific differentical expression association analysis: one known loci in *SREBF1* and one new loci in *FCRL6*. Further biological functional experiments are requried in more directly glucose-related tissues, e.g. pancreatic cells and liver, to unravel the mechanisms.

Online methods

Study population

The discovery samples consisted of 4,808 European individuals without diabetes from four non-overlapped cohorts, recruited by Rotterdam Study III-1 (RS III-1, n = 626), Rotterdam Study II-3 and Rotterdam Study III-2[45] (called as RS-BIOS, n = 705), Netherlands Twin Register[46,47] (NTR, n = 2,753) and UK adult Twin registry[48] (TwinsUK, n = 724). The replication sets contained up to 11,750 individuals from 11 independent cohorts from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE), including up to 6,818 individuals from European ancestry, 4,355 from African ancestry and 577 from Hispanic ancestry. We excluded individuals with known diabetes, those on anti-diabetic treatment or fasting glucose ≥ 7mmol/l. Local research ethics committees approved each study, and all participants gave informed consent to each original study. The details of the cohorts and the study design are shown in **Supplementary Note.**

Glycemic traits and covariates

Venous blood samples were obtained after an overnight fast in all discovery and replication cohorts. Details of fasting glucose and insulin measurements are shown in **Supplementary Note**. Body mass index (BMI) was calculated as weight over height squared (kg/m^2) based on clinical examinations. Smoking status was divided into current, former and never, based on questionnaires. White blood cell counts were quantified using standard laboratory techniques or predicted from methylation data using the standard Houseman method[49] (see **Supplementary Note** for each cohort).

DNA methylation quantification

The Illumina Human Methylation450 array was used in all discovery and replication cohorts to quantify genome-wide DNA methylation in blood samples. We obtained DNA methylation levels reported as β values, which represents the cellular average methylation level ranging from 0 (fully unmethylated) to 1 (fully methylated). Study-specific details regarding DNA methylation quantification, normalization and quality control procedures are provided in the **Supplementary Note** and **Supplementary Table 1**.

Epigenome-wide association analysis and replication

All statistical analyses were performed using *R* statistical software. Insulin was natural log transformed. In the discovery analysis, we first performed epigenome-wide association studies (EWAS) in each cohort separately. Linear regression analysis was used to test the association between glucose and insulin with each methylation site in the Rotterdam Study samples. Linear mixed models were used in NTR and TwinsUK accounting the

family structure. We fitted two models for each cohort: 1) the baseline model adjusting for age, sex, technical covariates (chip array number and position on the array), white blood cell counts (lymphocytes, monocytes, and granulocytes) and smoking status, and 2) a second model additionally adjusting for BMI. We removed probes that have evidence of multiple mapping or contain a genetic variant in the methylation site[50]. All cohort-specific EWAS results for each model were then meta-analysed using inverse variance-weighted fixed effects meta-analysis as implemented in the "metafor" R package[51]. In total, we meta-analysed 403,011 CpGs that passed quality control in all four discovery cohorts. The detail of the quality control for each cohort could be found in the **Supplementary Note**. The association was later corrected by the genomic control factor (λ) in each meta-EWAS[52]. We produced quantile-quantile (QQ) plots of the -$\log_{10}$ (*P*) to evaluate inflation in the test statistic (**Supplementary Figure 1**). A Bonferroni correction was used to correct for multiple testing and identify epigenome-wide significant results ($P < 1.27 \times 10^{-7}$). We did not correct the number of glycemic traits and models, as they are highly correlated and not independent. The genome coordinates were provided by Illumina (GRCh37/hg19). The correlation of the CpG sites located in the same gene was further checked in the overall RS III-1 and RS-BIOS samples by Pearson's correlation test (n = 1,544) to find the independent top CpGs. For the associations discovered in the meta-EWAS that have not been reported previously, we attempted replication in independent samples using the same traits and models as in the discovery analyses. Study-specific details of replication cohorts are provided in **Supplementary Table 1** and **Supplementary Note**. Results from each replication cohort were meta-analysed using the same methods as in the discovery analyses. Bonferroni P-value < 3.3×10^{-3} (0.05 corrected by 15 loci tested for associations) was considered significant.

Genomics of the differentially methylated sites and glycemic traits

We identified the genetic determinants of the significant CpG sites known or replicated through the current EWAS using the results of the *cis* and *trans* methylation quantitative trait loci (meQTLs) from European blood-based BIOS database[26] from the Biobanking and BioMolecular resources Research Infrastructure of the Netherlands (BBMRI-NL) which captured meQTLs, expression quantitative trait loci (eQTLs) and expression quantitative trait methylations (eQTMs) from genome-wide database of 3,841 Dutch blood samples (See resources of the database in **URLs**). All the reported single-nucleotide polymorphisms (SNPs) with P-value adjusted for false discovery rate (FDR) less than 0.05 in the database were treated as the target genetic variants in the present study. The SNPs were annotated based on the information in the previous study[26] or the nearest protein-coding gene list from SNPnexus[53,54] on GRCh37/hg19.

We explored the associations of these DNA methylation-related SNPs with type 2 diabetes (T2D) or related traits, i.e. fasting glucose, insulin, hemoglobin A1c (HbA1c), based on public genome-wide association study (GWAS) datasets in European ancestry[4-7,27]. We checked the effect direction consistency of the association between the SNPs, methylation sites and T2D or related traits. That is the direction of the association between SNP and T2D or related traits should be a combination of the directions of SNP with methylation sites and methylation sites with T2D or related traits. A multiple-testing correction was performed by Bonferroni adjustment (P-value < 1.8×10^{-3}, 0.05 corrected by the 29 genetic loci shown in **Supplementary Table 4).**

For the significant CpG sites known or replicated through EWAS, we attempted to evaluate the causality effect of CpGs on their significant traits, either fasting glucose or fasting insulin, using two-sample Mendelian Randomization (MR) approach as described in detail before by Dastani *et al*[30,55,56] based on the summary statistic GWAS results from BIOS database and the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) database[5,26] (**Supplementary Figure 2).** Briefly, we constructed a weighted genetic risk score for individual CpG on phenotype using independent SNPs as the instrument variables of the CpG, implemented in the R-package "*gtx*". The effect of each score on phenotype was calculated as

$$ahat = \frac{\sum(\omega_i \beta_i / s_i^2)}{\sum(\omega_i^2 / s_i^2)}$$

, where β_i is the effect of the CpG-increasing alleles on phenotype, $s_i$its corresponding standard error and ω_i the SNP effect on the respective CpG. Because the genetic variants might be close (*cis*) or far (*trans*) from the methylated site, we also performed MR test in the *cis* only SNPs if the CpG site has both *cis* and *trans* genetic markers. All SNPs were mapped to human genome build hg19. For each test (one CpG site with one trait), we extracted all the genetic markers of the CpG site in the fasting glucose or insulin GWAS from the MAGIC dataset (n = 96,496)[5] with their effect estimate and standard error on fasting glucose or insulin. Within the overlapped SNPs, we removed SNPs in potential linkage disequilibrium (LD, pairwise $R^2 \geq 0.05$) in 1-Mbp window based on the 1000 Genome imputed genotype dataset from the general population: Rotterdam Study I (RSI, n = 6,291)[45]. We managed to exclude the genetic loci which were genome-wide significantly associated with glycemic traits, but none of the genetic loci meet this exclusion criteria. The instrumental variables that explain more than 1% of variance in exposure (DNA methylation) were taken forward for MR test. The Bonferroni P-value threshold was used to correct for the 13 CpG sites available for MR (P-value < 3.8×10^{-3}).

Gene expression analyses

To explore whether the differential methylation sites were associated with differential expression in blood, we explored the European blood-based BIOS database for eQTMs[26]. The significantly associated gene expression probes were searched in the eQTL data in BIOS database[26]. We also investigated if the genetic variants associated with these gene expression probes in blood were also related to the DNA methylation sites with glycemia. Finally, we tested whether the expression of the genes that harbor the eQTMs was associated with T2D and related traits in glucose metabolism-related tissues (adipose subcutaneous, adipose visceral omentum, liver, whole blood, pancreas, and muscle skeletal) using *MetaXcan*[28,29,31] package. MetaXcan associates the expression of the genes with the phenotype by integrating functional data generated by large-scale efforts, e.g Genotype-Tissue Expression (GTEx)[32,33] with that of the GWAS of the trait. MetaXcan is trained on transcriptome models in 44 human tissues from GTEx and is able to estimate their tissue-specific effect on phenotypes from GWAS. For this study we used the GWAS studies of T2D[9], fasting glucose traits[5,6], fasting insulin[6], hemoglobin A1c (HbA1c)[57] and homeostatic model assessment-insulin resistance (HOMA-IR)[4]. We used the nominal P-value threshold (P-value = 0.05) as we had separate assumptions for each terminal pathway between gene expressions and phenotype. Further, we checked the effect direction consistency of the association between the methylation sites and fasting glucose or insulin with the combination of the associations between the methylation sites and gene expression and between the gene expression and T2D or related traits.

URLs. BIOS database, https://genenetwork.nl/biosqtlbrowser/; SNPnexus, http://snp-nexus.org/index.html; GWAS database of glycemic traits, https://www.magicinvestigators.org/; GWAS database of T2D, http://diagram-consortium.org/; MetaXcan, https://s3.amazonaws.com/imlab-open/Data/MetaXcan/results/metaxcan_results_database_v0.1.tar.gz. (available: 1st Jan, 2018)

Table1 Epigenome-wide association study (EWAS) results: novel differentially methylated sites associated with fasting glucose or insulin at an epigenome-wide significance level

Locus	CpG	Chr	Position	Regulatory feature	Trait (s)	Discovery phase (EA)				Replication phase (EA+AA+HS)			
						Model 1		Model 2		Model 1		Model 2	
						Z	P-value*	Z	P-value*	Z	P-value	Z	P-value
FCRL6	cg00936728	1	159772194	NA	Glucose	**-6.17**	**9.1×10^{-8}**	-5.71	1.9×10^{-7}	**-3.90**	**9.55×10^{-5}**	NP	NP
SLAMF1	cg18881723	1	160616870	Promoter associated	Glucose	**6.11**	**7.5×10^{-8}**	**6.94**	**3.4×10^{-10}**	2.67	7.66×10^{-3}	**3.23**	**1.2×10^{-3}**
1q25.3	cg13222915	1	184598594		Insulin	**-6.50**	**2.6×10^{-9}**	-4.61	4.1×10^{-6}	**-8.16**	**3.33×10^{-16}**	NP	NP
BRE	cg20657709	2	28509570	NA	Glucose	-5.46	2.7×10^{-6}	**-5.88**	**4.1×10^{-8}**	NP	NP	-2.10	0.036
LRPPRC	cg01913188	2	44223249	Promoter associated	Glucose	5.13	9.4×10^{-6}	**6.27**	**5.7×10^{-9}**	NP	NP	0.12	0.90
IRAK2	cg14527942	3	10276383	NA	Insulin	**6.97**	**3.4×10^{-10}**	**6.69**	**2.9×10^{-11}**	-0.70	0.48	-0.75	0.45
LETM1	cg13729116	4	1859262	Promoter associated	Insulin	**5.95**	**4.3×10^{-8}**	4.56	4.5×10^{-6}	**4.96**	**6.98×10^{-7}**	NP	NP
RBM20	cg15880704	10	112546110	NA	Insulin	**6.41**	**3.8×10^{-9}**	1.38 4.06	6.7×10^{-5}	**6.83**	**8.62×10^{-12}**	NP	NP
IRS2	cg25924746	13	110432935	Promoter associated	Insulin	**6.59**	**3.0×10^{-9}**	4.55	4.9×10^{-6}	**6.65**	**3.01×10^{-11}**	NP	NP
SPTB	cg07119168	14	65225253	NA	Glucose	-5.86	4.4×10^{-7}	**-5.82**	**4.9×10^{-8}**	NP	NP	-1.81	0.070
15q26.1	cg18247172	15	91370233	NA	Glucose	-5.25	4.9×10^{-6}	**-5.90**	**2.8×10^{-8}**	NP	NP	**-3.47**	**5.1×10^{-4}**
MAN2A2	cg20507228	15	91460071	Promoter associated (Cell type specific)	Insulin	**5.90**	**5.5×10^{-8}**	4.83	9.0×10^{-7}	**7.93**	**2.28×10^{-15}**	NP	NP
FAM92B	cg06709610	16	85143924	NA	Insulin	**6.35**	**6.5×10^{-9}**	**7.24**	**5.8×10^{-13}**	0.24	0.81	0.55	0.59
CD300A	cg08087047	17	72461209	NA	Glucose	-5.19	5.9×10^{-6}	**-5.80**	**1.1×10^{-7}**	NP	NP	-1.08	0.28
APOBEC3H	cg06229674	22	39492189	NA	Glucose	-5.40	1.8×10^{-6}	**-5.86**	**4.7×10^{-8}**	NP	NP	**-4.82**	**1.4×10^{-6}**

Genome-wide DNA methylation sites were tested for association with fasting glucose or fasting insulin in two models. Novel epigenome-wide significant (P-value < 1.27×10^{-7}) results in the discovery phase and the replication are shown. Model 1 adjusted for age, sex, technical covariates, white blood cell, and smoking status, accounting for family structure if needed in each cohort. Model 2 adjusted for BMI additionally. The significant associations in non-reported CpG sites were promoted for replication of the same models and traits. NP: Replication was not performed in the non-significant associated model or trait. Locus: the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Regulatory feature: the regulatory feature group of the CpGs from Illumina annotation. Chr: Chromosome. * Genomic controlled P-value. EA+AA+HA: European ancestry, African ancestry and Hispanic ancestry population. Z: effect estimate per standard error. **Bold print**: Significant results (P-value < 1.27×10^{-7} in the discovery phase, P-value < 3.3×10^{-3} in the replication phase). NA: Not available.

Supplementary Table 1 Characteristics of cohorts

COHORT	Discovery cohorts				Replication cohorts																
	RS III-1	RS-BIOS	NTR	Twins UK	ARIC	BLSA		CHS		FHS	GENOA	GOLDN	HyperGEN	inCHIANTI	KORA	WHI-BAA23			WHI-EMPC		
Ethnicity	European	European	European	European	African American	European	African American	European	African American	European	African American	European	African American	European	European	European	African American	Hispanic	European	African American	Hispanic
Fasting glucose (N)	626	705	2753	724	1875	402	142	160	147	2117	268	917	469	433	1488	836	503	319	425	951	258
Fasting insulin (N)	626	705	2685	724	NA	402	142	160	147	2157	267	915	466	421	1488	817	495	314	413	917	246
Fasting glucose (mmol/L)	5.3 (0.5)	5.4 (0.6)	5.2 (0.6)	5.0 (0.5)	5.7 (0.6)	5.3 (0.5)	5.3 (0.5)	5.5 (0.5)	5.5 (0.6)	5.6 (0.5)	5.3 (0.6)	5.48 (0.67)	5.37 (1.3)	4.9 (0.6)	5.3 (0.5)	5.28 (0.5)	5.2 (0.6)	5.2 (0.5)	5.2 (0.6)	5.2 (0.6)	5.2 (0.6)
Fasting insulin (pmol/L; mIU/L)	89.0 (44.3)	79.4 (44.4)	57.0 (36.1)	57.0 (53.6)	NA	58.8 (38.9)	78.7 (54.2)	76.3 (38.8)	91.4 (63.7)	4.07 (0.6)	62.4 (55.8)	13.7 (7.5)	9.4 (9.0)	72.8 (37.9)	45.8 (82.9)	5.3 (0.8)	5.79 (0.7)	5.7 (0.7)	90.2 (44.3)	75.2 (41.0)	88.3 (46.6)
Age (years)	59.8 (8.1)	67.5 (6.0)	36.3 (12.8)	58.1 (9.3)	56.1 (5.8)	69.0 (14.4)	64.3 (10.9)	76.2 (5.0)	73.2 (5.7)	65.7 (8.9)	60.7 (7.7)	48.0 (15.8)	51.0 (13.7)	62.1 (16.0)	60.2 (8.8)	68.4 (6.2)	62.6 (6.7)	62.4 (6.8)	62.1 (6.9)	64.5 (7.2)	61.4 (6.1)
BMI (kg/m2)	27.4 (4.5)	27.6 (4.1)	24.1 (3.7)	26.4 (4.6)	29.3 (6.0)	26.4 (4.4)	287 (4.9)	26.3 (4.6)	28.4 (4.7)	27.7 (4.9)	30.5 (6.4)	28.5 (5.6)	30.7 (7.0)	27.0 (3.9)	27.7 (4.5)	28.2 (5.4)	30.9 (6.4)	28.7 (4.9)	31.2 (6.0)	28.4 (5.7)	29.1 (4.8)
Female (%)	351 (56)	414 (59)	1817 (66)	684 (100)	1198 (64)	202 (50)	85 (60)	89 (56)	99 (67)	1204 (57)	191 (71)	488 (49)	365 (60)	234 (54)	783 (53)	836 (100)	503 (100)	319 (100)	425 (100)	951 (100)	258 (100)
Never Smoker (%)	187 (30)	244 (35)	1625 (59)	423 (62)	841 (45)	219 (54)	90 (63)	73 (46)	74 (50)	771 (37)	151 (56)	700 (70.3)	84 (17.8)	244 (56)	662 (44)	438 (52)	234 (467)	188 (59)	207 (49)	482 (51)	168 (65)
Former Smoker (%)	270 (43)	389 (55)	633 (23)	197 (30)	558 (30)	173 (43)	48 (34)	64 (40)	51 (35)	1180 (56)	69 (26)	209 (21)	207 (44.9)	105 (24)	607 (41)	381 (46)	258 (51)	123 (59)	176 (41)	402 (42)	73 (28)

NA: Not available. The unit of fasting insulin in FHS, GOLDN, HyperGEN, WHI-BAA23 are mIU/L; the units of fasting insulin in other cohorts are pmol/L.

Supplementary Table 2 Epigenome-wide association study (EWAS) results: known differentially methylated sites associated with fasting glucose or insulin at epigenome-wide significance level in the discovery phase

Locus	CpG	Chr	Position	Regulatory feature	Trait (s)	Model 1		Model 2		Previous evidence
						Z	P-value*	Z	P-value*	
DHCR24	cg17901584	1	55353706	Promoter associated (Cell type specific)	Insulin	**-6.19**	**2.3×10^{-8}**	-3.58	3.7×10^{-4}	BMI[13]; Insulin[13]; HbA1c[13]; Incident T2D[13]
4p15.33	cg10438589	4	14531493	NA	Insulin	**6.25**	**2.4×10^{-8}**	3.56	5.0×10^{-4}	BMI[13]; Insulin[13]; Incident T2D[13]
SLC7A11	cg06690548	4	139162808	NA	Glucose	**-7.70**	**7.6×10^{-10}**	**-5.85**	**7.8×10^{-8}**	BMI[13]; Glucose[13]; Insulin[13]; Incident T2D[13]
					Insulin	**-6.21**	**2.4×10^{-8}**	-3.68	2.8×10^{-4}	
RNF145	cg26403843	5	158634085	NA	Insulin	**6.29**	**8.5×10^{-9}**	4.60	7.7×10^{-6}	BMI[13,17,20]; Insulin[13]; Incident T2D[13]
CPT1A	cg00574958	11	68607622	NA	Glucose	**-7.63**	**2.9×10^{-11}**	-5.56	2.5×10^{-7}	BMI[13,17,18,20]; Glucose[13,22]; Insulin[13]; HbA1c[13]; Incident T2D[13]; Prevalent T2D[17]
					Insulin	**-8.00**	**3.5×10^{-13}**	-4.11	3.5×10^{-5}	
CPT1A	cg17058475	11	68607737	NA	Insulin	**-6.36**	**7.4×10^{-9}**	-4.16	3.4×10^{-5}	BMI[13]; Glucose[13]; Insulin[13]; HbA1c[13]; Incident T2D[13]
ASAM	cg26894079	11	122954435	NA	Insulin	**-5.83**	**7.5×10^{-8}**	-2.95	3.3×10^{-3}	BMI[13]; Insulin[13]; Incident T2D[13]
KDM2B	cg13708645	12	121974305	Promoter associated	Insulin	**6.00**	**1.1×10^{-7}**	3.28	8.8×10^{-4}	BMI[17,20]
MYO5C	cg06192883	15	52554171	Unclassified	Insulin	**7.96**	**6.4×10^{-13}**	4.57	4.4×10^{-6}	BMI[13,17,20]; Insulin[13]; Incident T2D[13]
SREBF1	cg11024682	17	17730094	Unclassified (Cell type specific)	Glucose	**6.45**	**2.7×10^{-8} 6.7**	4.34	6.4×10^{-5}	BMI[13,20,24]; Glucose[13,22]; Insulin[13]; HbA1c[13]; Incident T2D[13,19]; Prevalent T2D[17]
					Insulin	**6.27**	**$\times 10^{-9}$**	2.69	8.6×10^{-3}	
TMEM49	cg24174557	17	57903544	NA	Insulin	**-7.57**	**8.8×10^{-12}**	-4.00	6.7×10^{-5}	BMI[13,17]; Insulin [13]; Incident T2D[13]
ABCG1	cg27243685	21	43642366	NA	Insulin	**7.55**	**5.9×10^{-12}**	5.10	4.5×10^{-7}	BMI[13,17,20,24,25]; Glucose[13]; Insulin[13]; HbA1c[13]; Incident T2D[13]
ABCG1	cg06500161	21	43656587	NA	Insulin	**10.16**	**$< 2.2 \times 10^{-16}$**	**6.68**	**5.0×10^{-11}**	BMI[13,17,20,21]; Glucose[13,22]; Insulin[12,13,22]; HbA1c[13]; Incident T2D[13,19]; Prevalent T2D[17,23]; 2h glucose[22]; HOMA-IR[12]

Genome-wide DNA methylation sites were tested for association with fasting glucose or fasting insulin in two models. Previously reported epigenome-wide significant (P-value < 1.27×10^{-7}) results in the discovery phase and the previous evidence from the EWAS with T2D or related traits in the same CpG sites are shown. Model 1 adjusted for age, sex, technical covariates, white blood cell, and smoking status, accounting for family structure if needed in each cohort. Model 2 adjusted for BMI additionally. Locus: the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Regulatory feature: the regulatory feature group of the CpGs from Illumina annotation. Chr: Chromosome. * Genomic controlled P-value. Z: effect estimate per standard error. **Bold print**: Epigenome-wide significant results (P-value < 1.27×10^{-7}). NA: Not available.

Supplementary Table 3 Epigenome-wide association study (EWAS) results: replication of newly discovered differentially methylated sites in different ancestry populations

Locus	CpG	Chr	Position	Trait (s)	Replication in EA				Replication in AA				Replication in HA			
					Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
					Z	P-value	Z	P-value	Z	P-value	Z	P-value	Z	P-value	Z	P-value
FCRL6	cg00936728	1	159772194	Glucose	-2.48	0.013	NP	NP	-2.90	3.8×10^{-3}	NP	NP	-1.84	0.066	NP	NP
SLAMF1	cg18881723	1	160616870	Glucose	2.68	7.3×10^{-3}	**3.07**	**2.1×10^{-3}**	1.14	0.25	1.59	0.11	-0.56	0.58	-0.49	0.63
1q25.3	cg13222915	1	184598594	Insulin	**-6.85**	**7.6×10^{-12}**	NP	NP	**-5.26**	**1.4×10^{-7}**	NP	NP	0.10	0.092	NP	NP
BRE	cg20657709	2	28509570	Glucose	NP	NP	-0.93	0.35	NP	NP	-2.43	0.015	NP	NP	0.23	0.82
LRPPRC	cg01913188	2	44223249	Glucose	NP	NP	0.08	0.94	NP	NP	-0.27	0.79	NP	NP	0.99	0.32
IRAK2	cg14527942	3	10276383	Insulin	-1.06	0.29	-1.15	0.25	0.63	0.53	0.45	0.66	0.21	0.83	0.65	0.51
LETM1	cg13729116	4	1859262	Insulin	**4.63**	**3.6×10^{-6}**	NP	NP	1.69	0.092	NP	NP	NP	NP	0.62	0.54
RBM20	cg15880704	10	112546110	Insulin	**6.83**	**8.6×10^{-12}**	NP	NP	1.22	0.23	NP	NP	NP	NP	0.55	0.59
IRS2	cg25924746	13	110432935	Insulin	**6.20**	**5.7×10^{-10}**	NP	NP	1.85	0.064	NP	NP	NP	NP	1.65	0.10
SPTB	cg07119168	14	65225253	Glucose	NP	NP	-0.57	0.57	NP	NP	-2.23	0.026	NP	NP	-0.97	0.33
15q26.1	cg18247172	15	91370233	Glucose	NP	NP	-1.69	0.092	NP	NP	**-3.66**	**2.6×10^{-4}**	NP	NP	-1.18	0.24
MAN2A2	cg20507228	15	91460071	Insulin	**7.59**	**3.2×10^{-14}**	NP	NP	1.59	0.11	NP	NP	2.75	6.0×10^{-3}	NP	NP
FAM92B	cg06709610	16	85143924	Insulin	0.82	0.41	0.83	0.41	-1.93	0.053	-1.37	0.17	1.13	0.26	1.65	0.10
CD300A	cg08087047	17	72461209	Glucose	NP	NP	-0.95	0.34	NP	NP	-0.21	0.83	NP	NP	-0.96	0.34
APOBEC3H	cg06229674	22	39492189	Glucose	NP	NP	**-3.88**	**1.0×10^{-4}**	NP	NP	-2.70	6.9×10^{-3}	NP	NP	-0.99	0.32

Replication of the epigenome-wide significant (P-value < 1.27×10^{-7}) CpGs with fasting glucose or insulin stratified by different ancestry populations. Model 1 adjusted for age, sex, technical covariates, white blood cell, and smoking status, accounting for family structure if needed in each cohort. Model 2 adjusted for BMI additionally. EA:

European ancestry (n = 6,778 for fasting glucose; n = 6,773 for fasting insulin) AA: African ancestry (n = 4,355 for fasting glucose; n = 2,434 for fasting insulin). HA: Hispanic ancestry (n = 577 for fasting glucose; n = 560 for fasting insulin). NP: Replication was not performed in the non-significant associated model or trait from the discovery phase. Locus: the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Chr: Chromosome. Z: effect estimate per standard error. **Bold print**: Bonferroni significant results (P-value < 3.3×10^{-3}).

Supplementary Table 4 Methylation quantitative trait loci (meQTLs) for known or new replicated CpG sites

Locus (CpG)	CpG	Variant	Chr	Position	Locus (meQTL)	Type (meQTL)	MAF	Alleles	EA	Z	P-value	*Cis*/*Trans*
DHCR24	cg17901584	rs7412	19	45412079	*APOE*	Protein coding	0.08	C/T	T	6.39	1.7×10^{-10}	*Trans*
DHCR24	cg17901584	rs7701414	5	131585958	*P4HA2*	Protein coding	0.22	A/G	G	5.26	1.4×10^{-7}	*Trans*
DHCR24	cg17901584	rs174550	11	61571478	*FADS1/FADS2*	Protein coding	0.30	T/C	C	-6.72	1.8×10^{-11}	*Trans*
DHCR24	cg17901584	rs735665	11	123361397	*GRAMD1B (Nearest)*	Protein coding	0.10	G/A	A	10.59	3.4×10^{-26}	*Trans*
DHCR24	cg17901584	rs687565	1	55364663	*TMEM61 (Nearest)*	Protein coding	0.43	C/A	C	11.95	6.2×10^{-33}	*Cis*
FCRL6	cg00936728	rs6657365	1	159782549	*FCRL6*	Protein coding	0.23	C/G	G	5.04	4.6×10^{-7}	*Cis*
FCRL6	cg00936728	rs2523946	6	29941943	*HCG9 (Nearest)*	lincRNA	0.49	C/T	T	5.33	9.8×10^{-8}	*Trans*
SLAMF1	cg18881723	rs3129055	6	29670261	*ZFP57 (Nearest)*	Protein coding	0.29	A/G	G	6.37	1.9×10^{-10}	*Trans*
SLAMF1	cg18881723	rs11265461	1	160630143	*SLAMF1 (Nearest)*	Protein coding	0.36	C/T	C	8.39	4.7×10^{-17}	*Cis*
1q25.3	cg13222915	rs72737737	1	184598732	*C1orf21*	Protein coding	0.06	A/G	G	5.68	1.4×10^{-8}	*Cis*
4p15.33	cg10438589	rs16890352	4	14385522	*AC006296.1*	lincRNA	0.16	A/G	G	-11.64	2.6×10^{-31}	*Cis*
RNF145	cg26403843	rs7529925	1	199007208	*RP11-16L9.4*	lincRNA	0.21	T/C	C	5.68	1.3×10^{-8}	*Trans*
RNF145	cg26403843	rs7732603	5	158614357	*RNF145*	Protein coding	0.48	A/C	C	35.61	8.6×10^{-278}	*Cis*
RBM20	cg15880704	rs7906643	10	112545494	*RBM20*	Protein coding	0.09	C/T	T	-29.45	1.4×10^{-190}	*Cis*
CPT1A	cg00574958	rs964184	11	116648917	*ZNF259*	Protein coding	0.22	C/G	G	-5.54	3.0×10^{-8}	*Trans*
ASAM	cg26894079	rs34817879	11	123023729	*CLMP*	Protein coding	0.12	T/G	G	-4.67	3.0×10^{-6}	*Cis*
KDM2B	cg13708645	rs60370741	12	121966676	*KDM2B*	Protein coding	0.30	T/C	C	37.16	2.6×10^{-302}	*Cis*
IRS2	cg25924746	rs9521528	13	110504805	*IRS2 (Nearest)*	Protein coding	0.44	T/A	T	-13.64	2.4×10^{-42}	*Cis*

IRS2	cg25924746	rs7984800	13	110671951	*LINC00396 (Nearest)*	lincRNA	0.26	A/G	G	3.88	1.0×10^{-4}	*Cis*
MYO5C	cg06192883	rs1047891	2	211540507	*CPS1*	Protein coding	0.29	A/C	A	-6.01	1.8×10^{-9}	*Trans*
MYO5C	cg06192883	rs71472932	15	52541976	*MYO5C*	Protein coding	0.11	G/A	A	4.22	2.4×10^{-5}	*Cis*
15q26.1	cg18247172	rs404623	15	91367271	*BLM (Nearest)*	Protein coding	0.50	G/C	G	-11.70	1.3×10^{-31}	*Cis*
15q26.1	cg18247172	rs3129055	6	29670261	*ZFP57 (Nearest)*	Protein coding	0.29	A/G	G	-5.22	1.8×10^{-7}	*Trans*
15q26.1	cg18247172	rs4324798	6	28776117	*AL662890.3*	miRNA	0.04	G/A	A	-5.48	4.3×10^{-8}	*Trans*
MAN2A2	cg20507228	rs9374080	6	109616420	*CCDC162P*	pseudogene	0.27	C/T	C	-5.31	1.1×10^{-7}	*Trans*
MAN2A2	cg20507228	rs35831960	15	91466262	*MAN2A2 (Nearest)*	Protein coding	0.17	C/T	T	8.14	4.1×10^{-16}	*Cis*
SREBF1	cg11024682	rs7701414	5	131585958	*P4HA2*	Protein coding	0.22	A/G	G	5.21	1.9×10^{-7}	*Trans*
SREBF1	cg11024682	rs7529925	1	199007208	*RP11-16L9.4*	lincRNA	0.21	T/C	C	-5.15	2.6×10^{-7}	*Trans*
SREBF1	cg11024682	rs6502629	17	17869642	*TOM1L2*	Protein coding	0.22	G/A	G	9.97	2.1×10^{-23}	*Cis*
TMEM49	cg24174557	rs3774937	4	103434253	*NFKB1*	Protein coding	0.25	C/T	C	-7.75	9.3×10^{-15}	*Trans*
ABCG1	cg06500161	rs225443	21	43658206	*ABCG1*	Protein coding	0.40	G/A	A	-7.16	8.1×10^{-13}	*Cis*
APOBEC3H	cg06229674	rs28583464	22	39486593	*APOBEC3G (Nearest)*	Protein coding	0.11	T/C	C	8.31	9.8×10^{-17}	*Cis*

Based on the European blood-based BIOS database (n = 3,841),[26] the meQTL information of known or new replicated CpG sites are shown. Locus (CpG): the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Locus (meQTL): the located or nearest protein-coding gene of the meQTL from UCSC annotation. Type (meQTL): the gene type of the meQTL. Chr: chromosome. MAF: minor allele frequency. EA: effect allele. Z: effect estimate per standard error.

Supplementary Table 5 Common genetic determinants of glycemia related methylation sites and T2D or related traits in blood

Variant	Locus (meQTL)	Type (meQTL)	Chr	Position	MAF	EA	Association with CpG				Association with T2D or related traits[†]		
							CpG	Locus (CpG)	Z	P-value	Trait	Z	P-value
rs6701489	*TMEM61* *(Nearest)*	Protein coding	1	55358459	0.07	T	cg17901584	*DHCR24*	4.82	1.4×10^{-6}	FG[6]	-3.43	8.5×10^{-4}
rs6896438	*RNF145* *(Nearest)*	Protein coding	5	158547876	0.36	C	cg26403843	*RNF145*	6.15	8.0×10^{-10}	FI[5]	3.82	1.4×10^{-4}
rs10849885	*KDM2B*	Protein coding	12	121881848	0.32	A	cg13708645	*KDM2B*	29.33	4.2×10^{-189}	FG[6]	4.17	2.2×10^{-5}
rs9374080	*CCDC162P*	pseudogene	6	109616420	0.27	C	cg20507228	*MAN2A2*	-5.31	1.1×10^{-7}	HbA1c[58]	-5.11	2.0×10^{-7}
rs3818717	*RAI1*	Protein coding	17	17707105	0.06	T	cg11024682	*SREBF1*	8.93	4.1×10^{-19}	T2D[27]	1.08	4.9×10^{-4}
rs7529925	*RP11-16L9.4*	lincRNA	1	199007208	0.21	C	cg11024682	*SREBF1*	-5.15	2.6×10^{-7}	HbA1c[58]	-3.60	2.5×10^{-4}
rs16960744	*TOM1L2*	Protein coding	17	17755259	0.37	A	cg11024682	*SREBF1*	4.87	1.1×10^{-6}	HbA1c[58]	3.11	1.5×10^{-3}

The common genetic determinants of glycemia related methylation sites and T2D or related traits are shown. Chr: chromosome. Locus (meQTL): the located or nearest protein-coding gene of the meQTL from UCSC annotation. Type (meQTL): the gene type of the meQTL. MAF: minor allele frequency. EA: effect allele. Locus (CpG): the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Z: effect estimate per standard error. FG: fasting glucose. FI: fasting insulin. Data sources of associations: 1) association with CpG was from the current discovery phase (n = 4,808), 2) associations with FG (n = 133,010), FI (n = 96,496), T2D (case/control: n = 81,412/370,832) and HbA1c (n = 159,940) were from the MAGIC and DIAGRAM GWAS database[5,6,27,58]

Supplementary Table 6 Mendelian randomization (MR) results

Exposure	Locus (CpG)	Trait	No. of SNPs	R^2 (%)	Effect	SE	P-value	Heterogeneity P-value	Type of SNPs	SNPs list	Z (exposure)	Z (outcome)
cg00574958	*CPT1A*	NP	1	0.79	NP	NP	NP	NP	All-*Trans*	rs964184	-5.54	0.42
cg06192883	*MYO5C*	NP	1	0.91	NP	NP	NP	NP	All-*Trans*	rs715	-5.94	1.77
cg06229674	*APOBEC3H*	FG	1	1.71	0.09	0.12	0.48	NA	All-*Cis*	rs6001423	8.17	0.71
cg06500161	*ABCG1*	FI	2	2.30	0.09	0.11	0.42	0.27	All-*Cis*	rs225443;rs225391	-7.16;6.18	-1.33;-0.31
cg10438589	*4p15.33*	FI	4	7.15	-0.03	0.06	0.60	0.50	All-*Cis*	rs16890358;rs9291625;rs13131008;rs10488977	-11.63;-7.83;6.91;6.06	-0.3;1.23;0.43;-0.94
cg11024682	*SREBF1*	FG	2	4.65	-0.06	0.07	0.44	0.44	All-*Cis*	rs8070432;rs6502629	9.12;9.97	0.05;-1.09
cg11024682	*SREBF1*	FI	2	4.65	0.02	0.07	0.81	0.84	All-*Cis*	rs8070432;rs6502629	9.12;9.97	0.02;0.31
cg13708645	*KDM2B*	FI	3	30.09	0.04	0.03	0.11	0.22	All-*Cis*	rs28604990;rs11065536;rs3935332	37.01;-9.78;7.28	1.76;1.13;1.06
cg15880704	*RBM20*	FI	5	23.38	0.06	0.03	0.04	0.91	All-*Cis*	rs7906643;rs11195272;rs4918591;rs4918537;rs10509930	-29.45;-7.86;7.14;-6.72;5.92	-1.83;-1.05;-0.33;-0.6;0.56
cg17901584	*DHCR24*	FI	4	8.33	-0.07	0.06	0.19	0.07	All-*CisTrans*	rs681123;rs735665;rs174546;rs445925	11.86;10.59;-6.75;5.5	-0.47;-1.33;2;1.69
cg17901584	*DHCR24*	FI	1	3.53	-0.04	0.08	0.64	NA	Sub-*Cis*	rs681123	11.86	-0.47
cg18247172	*15q26.1*	FG	3	5.82	-0.05	0.07	0.47	0.30	All-*CisTrans*	rs8038275;rs2518968;rs4324798	11.55;-8.12;-5.48	0.27;0.71;1.53
cg18247172	*15q26.1*	FG	2	5.04	-0.01	0.07	0.85	0.46	Sub-*Cis*	rs8038275;rs2518968	11.55;-8.12	0.27;0.71
cg18881723	*SLAMF1*	FG	2	2.85	0.19	0.09	0.05	0.59	All-*CisTrans*	rs11265461;rs3129055	8.39;6.37	1.89;0.76
cg18881723	*SLAMF1*	FG	1	1.80	0.23	0.12	0.06	NA	Sub-*Cis*	rs11265461	8.39	1.89

cg20507228	*MAN2A2*	FI	1	1.69	-0.07	0.12	0.59	NA	All-*Cis*	rs1266482	8.12	-0.55
cg24174557	*TMEM49*	FI	1	1.54	-0.07	0.13	0.60	NA	All-*Trans*	rs3774937	-7.75	0.53
cg25924746	*IRS2*	FI	2	5.89	0.05	0.07	0.48	0.85	All-*Cis*	rs9521528;rs11842277	-13.64;-7.01	-0.55;-0.49
cg26403843	*RNF145*	FI	8	42.56	0.01	0.02	0.60	0.03	All-*CisTrans*	rs6556405;rs3846687;rs12188300;rs6890049;rs4244439;rs2043269;rs17056747;rs7529925	39.82;-16.71;-10.3;8.38;-7.39;-7.1;-5.55;5.68	-0.3;-2.66;0.86;0.31;0.09;1.97;-0.28;2.03
cg26403843	*RNF145*	FI	7	41.73	0.01	0.02	0.78	0.07	Sub-*Cis*	rs6556405;rs3846687;rs12188300;rs6890049;rs4244439;rs2043269;rs17056747	39.82;-16.71;-10.3;8.38;-7.39;-7.1;-5.55	-0.3;-2.66;0.86;0.31;0.09;1.97;-0.28

Two-sample MR approach was performed to check the effect of known or replicated CpG sites on their significant traits, either fasting glucose or fasting insulin. We also performed MR test in the *cis*-only SNPs if the CpG site has both *cis* and *trans* genetic markers. Locus (CpG): the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. R2 (%): the percentage of explained variance in the exposure by genetic risk score. Effect/SE/P-value: The effect estimate / standard error / P-value of genetic risk score of the exposure on the outcome (MR results). Heterogeneity P-value: The P-value of the heterogeneity test among the SNPs. Type of SNPs: type of SNPs included in the genetic risk score: 1) All-*CisTrans*: all the genetic markers (included *cis* and *trans* SNPs) ; 2) All-*Cis*: only *cis* genetic markers available; 3) All-*Trans*: only *trans* genetic markers available; 4) Sub-*Cis*: the sub-analysis with the genetic markers in *cis*-only. Z (exposure): the effect estimate per standard error of the SNP on exposure (CpG) from exposure GWAS result; Z (outcome): the effect estimate per standard error of the SNPs on outcome (fasting glucose or insulin) from outcome GWAS result. NP: the genetic risk score has R^2 less than 1%, and the MR was not performed. FG: fasting glucose. FI: fasting insulin. NA: Not available.

Supplementary Table 7 Blood-based expression quantitative trait methylations (eQTMs): association between gene expression and the glycemia related methylation sites

Locus (CpG)	CpG	Probe	Probe-Chr	Probe-Pos	Gene expression	Z	P-value	*Cis*/*Trans*
FCRL6	cg00936728	ENSG00000181036	1	159770301	*FCRL6*	-13.09	4.0×10^{-39}	*Cis*
SLAMF1	cg18881723	ENSG00000026751	1	160709037	*SLAMF7*	5.84	5.4×10^{-9}	*Cis*
SLAMF1	cg18881723	ENSG00000122223	1	160832692	*CD244*	4.68	2.9×10^{-6}	*Cis*
SLAMF1	cg18881723	ENSG00000117090	1	160617085	*SLAMF1*	-4.10	4.1×10^{-5}	*Cis*
CPT1A	cg00574958	ENSG00000110090	11	68611878	*CPT1A*	-9.22	3.1×10^{-20}	*Cis*
SREBF1	cg11024682	ENSG00000072310	17	17740325	*SREBF1*	-7.84	4.5×10^{-15}	*Cis*
ABCG1	cg06500161	ENSG00000160179	21	43619799	*ABCG1*	-12.78	2.2×10^{-37}	*Cis*

The gene expressions associated with the glycemia related methylation sites are shown based on the European blood-based BIOS database (n = 3,841) [26]. Locus (CpG): the cytogenetic location or the gene symbol of the CpGs from Illumina annotation. Probe: The probe of the gene expression. Chr: chromosome. Pos: position. Z: effect estimate per standard error.

Supplementary Table 8 Common genetic determinants of glycemia related DNA methylation (methylation quantitative trait loci, meQTL) and gene expression (expression quantitative trait loci, eQTL) in blood.

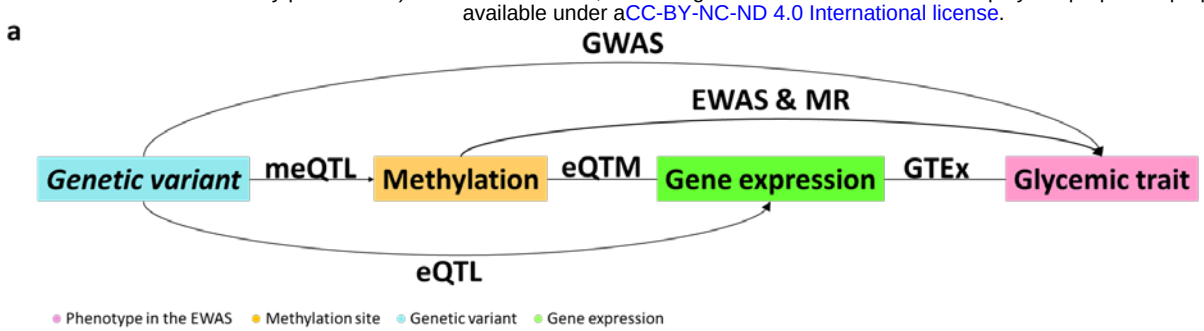
Variant	Chr	Position	Locus (eQTL)	Type (eQTL)	MAF	EA	Association with CpG					Association with gene expression			
							CpG	Locus (CpG)	Z	P-value	*Cis*/*Trans*	Gene expression	Z	P-value	*Cis*/*Trans*
rs11265282	1	159774408	*FCRL6*	Protein coding	0.26	C	cg00936728	*FCRL6*	4.17	3.0×10^{-5}	*Cis*	*FCRL6*	-6.73	1.7×10^{-11}	*Cis*
rs1577544	1	160630974	*SLAMF1* *(Nearest)*	Protein coding	0.39	T	cg18881723	*SLAMF1*	-5.45	5.1×10^{-8}	*Cis*	*SLAMF1*	-6.40	1.6×10^{-10}	*Cis*
rs6502629	17	17869642	*TOM1L2*	Protein coding	0.22	G	cg11024682	*SREBF1*	9.97	2.1×10^{-23}	*Cis*	*SREBF1*	-17.93	7.2×10^{-72}	*Cis*

The common genetic determinants of glycemia related methylation sites and gene expression in the European blood-based BIOS database (n = 3,841) [26] are shown. Locus (eQTL): the located or nearest protein-coding gene of the eQTL. Type (eQTL): the gene type of the eQTL. Chr: chromosome. MAF: minor allele frequency. EA: effect allele. Z: effect estimate per standard error. Locus (CpG): the cytogenetic location or the gene symbol of the CpGs from Illumina annotation.

Supplementary Table 9 Association between the gene expression level in the glucose metabolism-related tissue and the T2D or related traits based on the Genotype-Tissue Expression (GTEx) project

Gene expression	Trait	Tissue	Z	P-value
FCRL6	T2D[9]	Liver	2.14	0.032
SREBF1	T2D[9]	Whole blood	-2.40	0.016
SREBF1	HbA1c[4]	Whole blood	-3.26	1.1×10^{-3}

The significant associations between the gene expression level in the glucose metabolism-related tissue and the T2D or related traits are shown based on the tissue-specific Genotype-Tissue Expression (GTEx) project[28,29]. It was explored in six glucose related tissues, i.e. adipose subcutaneous, adipose visceral omentum, liver, whole blood, pancreas, and muscle skeletal, and five T2D or related traits, i.e. T2D[9], fasting glucose[5,6], fasting insulin[6], HbA1c[57], and HOMA-IR[4]. Z: effect estimate per standard error.



Abbreviation	Full name	Data source and sample size	Tissue
meQTL	Eethylation quantitative trait loci	Summary statistics from BIOS database (n=3,814)	Blood
eQTL	Expression quantitative trait loci	Summary statistics from BIOS database (n=3,814)	Blood
eQTM	Expression quantitative trait methylation	Summary statistics from BIOS database (n=3,814)	Blood
EWAS	Epigenome-wide association study	Current EWAS results (discovery n=4,808, replicated n=11,750)	Blood
MR	Mendelian randomization	Two-sample MR based on BIOS database (n=3,814) and MAGIC (n=96,496)	Blood
GWAS	Genome-wide association study	Summary statistics from MAGIC and DIAGRAM (n=96,496 ~ 452,244)	Blood
GTEx	Genotype-Tissue Expression project	Summary statistic from GTEx and MAGIC/DIAGRAM (tissues n=153 ~ 491)	Glucose related tissues

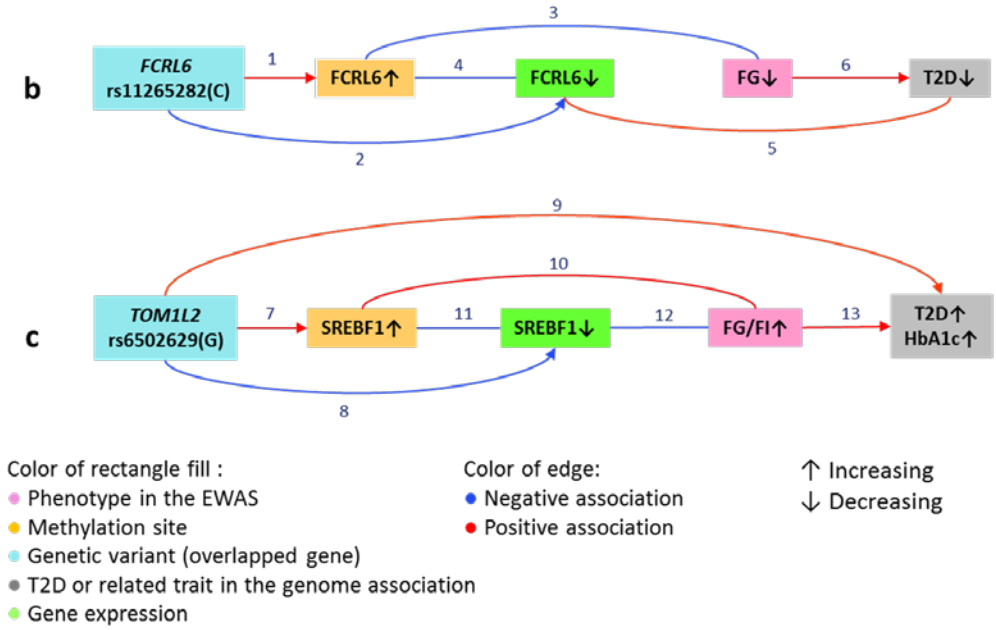
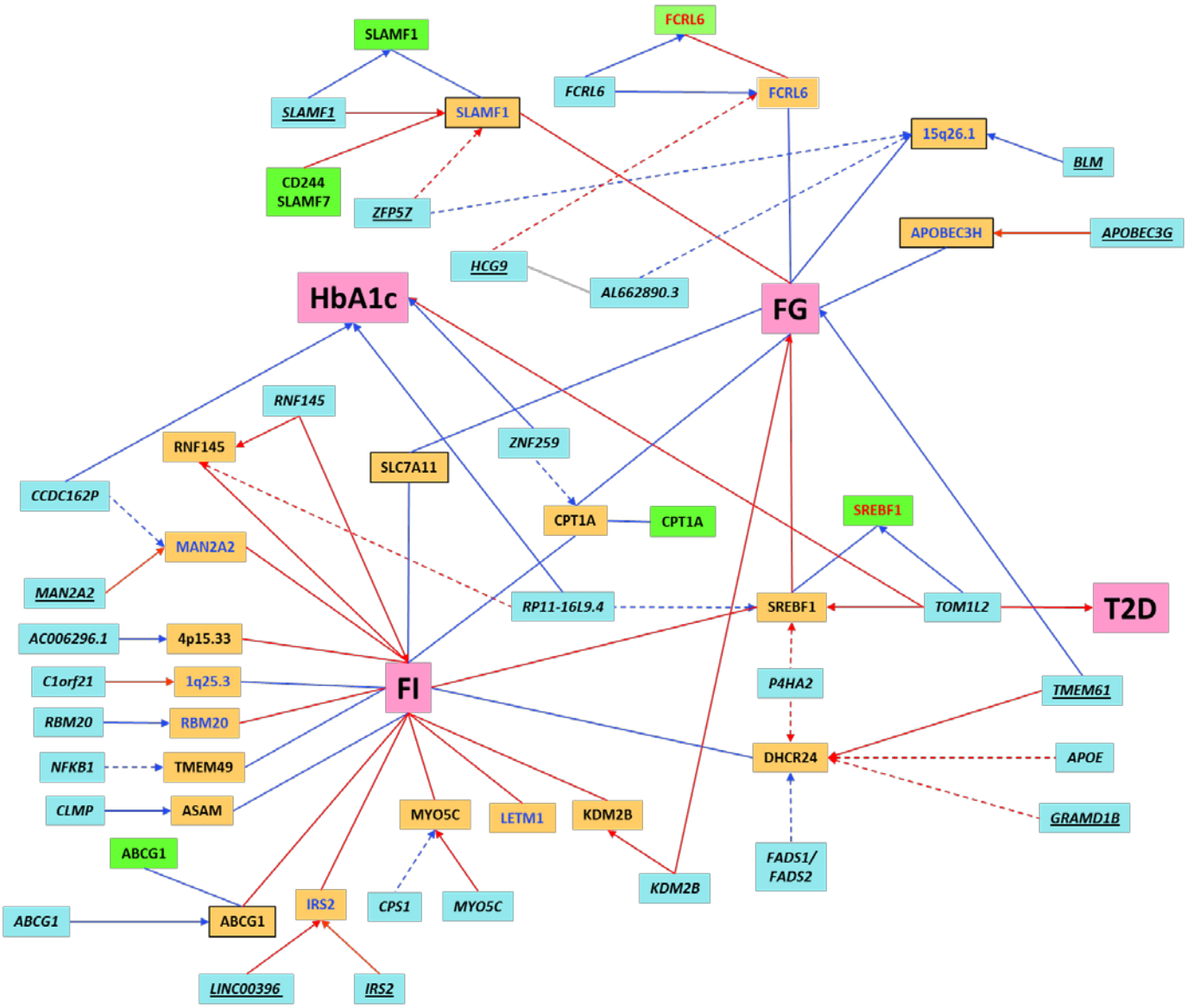


Fig.	NO.	Explanation	Z	P-value
b	1	meQTL	4.17	3.0×10^{-5}
	2	eQTL	-6.73	1.7×10^{-11}
	3	EWAS	-6.17	9.1×10^{-8}
	4	eQTM	-13.09	4.0×10^{-39}
	5	GTEx	2.14	0.032
	6	General Knowledge		
c	7	meQTL	9.97	2.1×10^{-23}
	8	eQTL	-17.93	7.2×10^{-72}
	9	GWAS	3.11	1.5×10^{-3}
	10	EWAS with FG	6.45	2.7×10^{-8}
	10	EWAS with FI	6.27	6.7×10^{-9}
	11	eQTM	-7.84	4.5×10^{-15}
	12	GTEx	-3.26	1.1×10^{-3}
	13	General Knowledge		

Figure 1 Overview of the cross-omics analysis and examples

Cascading associations cross multiple-omics-based on different data sources were integrated in the network figures. The assumption is genetic variants drive DNA methylation which subsequently regulates gene expression and then glycemic traits. FG: fasting glucose. FI: fasting insulin. T2D: type 2 diabetes



Color of rectangle fill :
- Phenotype in the EWAS
- Methylation site
- Genetic variant (overlapped or nearest gene)
- T2D or related trait in the genome association
- Gene expression

Color of rectangle outline in methylation site:
- BMI-independent methylation site

No outline: BMI-dependent methylation site

Color of edge:
- Negative association
- Positive association
- Association with direction not clear
- Linkage disequilibrium ($R^2 > 0.2$)

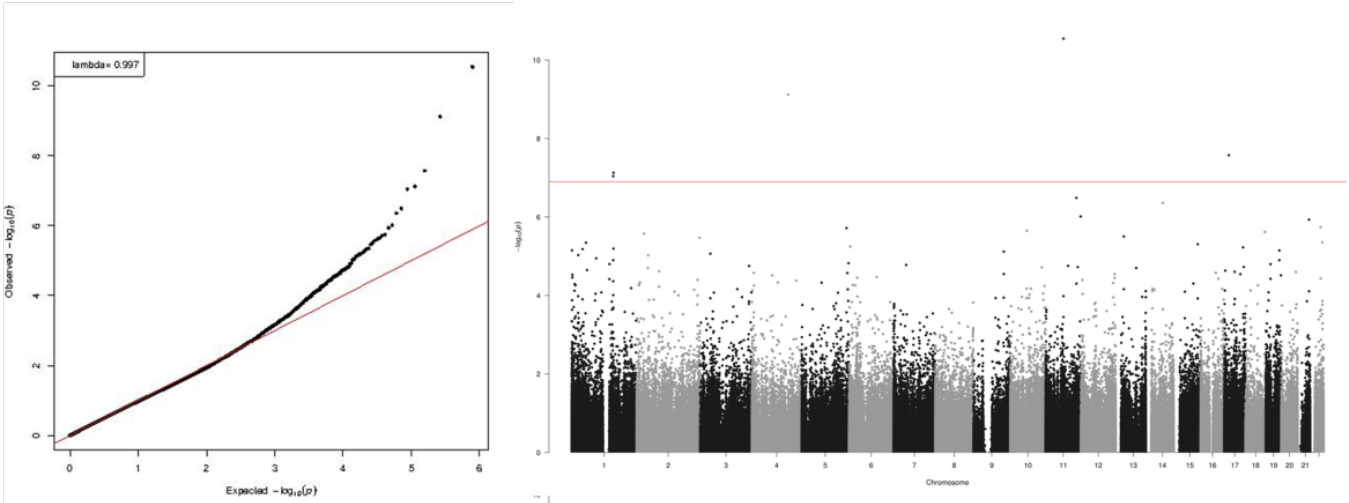
Color of letter:
- Novel methylation site and replicated successfully
- Gene expression correlated with T2D or related traits in the glucose metabolism-related tissue
- Others (gene, phenotype, other methylation site or gene expression)

Dash of edge:
-- meQTL or eQTL in trans
–> Casual association
– Others (LD, EWAS)

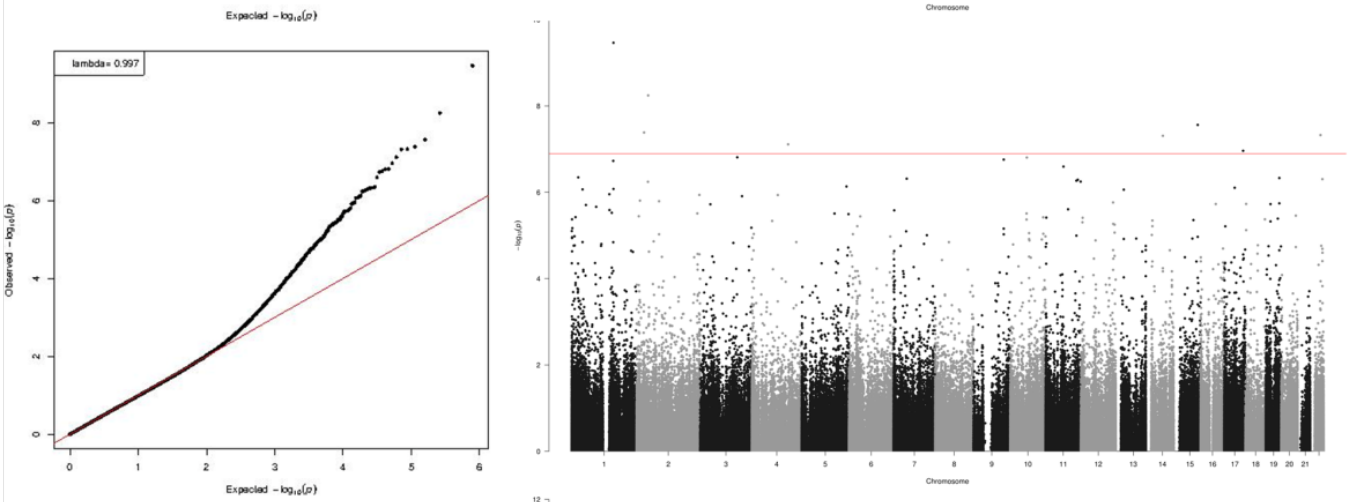
Figure 2 Associations between genetic variants, DNA methylation sites, gene expressions and fasting glucose, insulin and related traits based on the integration of cascading associations in Figure 1a

The effect allele is standardized across all associations. Only the significant associations which passed the specific P-value threshold in each association step were shown in the figure. FG: fasting glucose. FI: fasting insulin. T2D: type 2 diabetes.

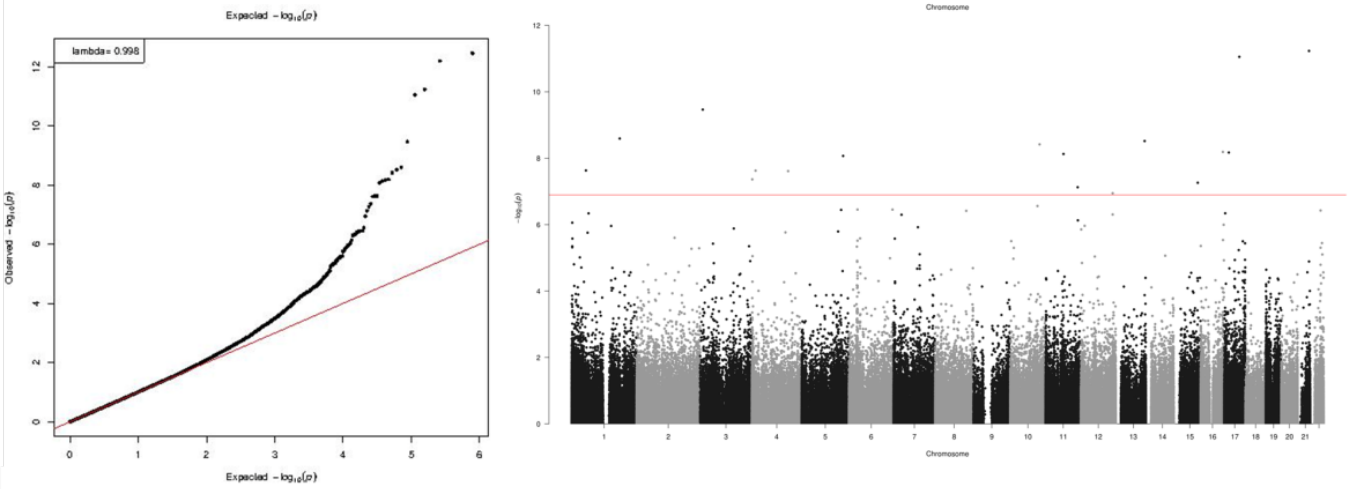
A



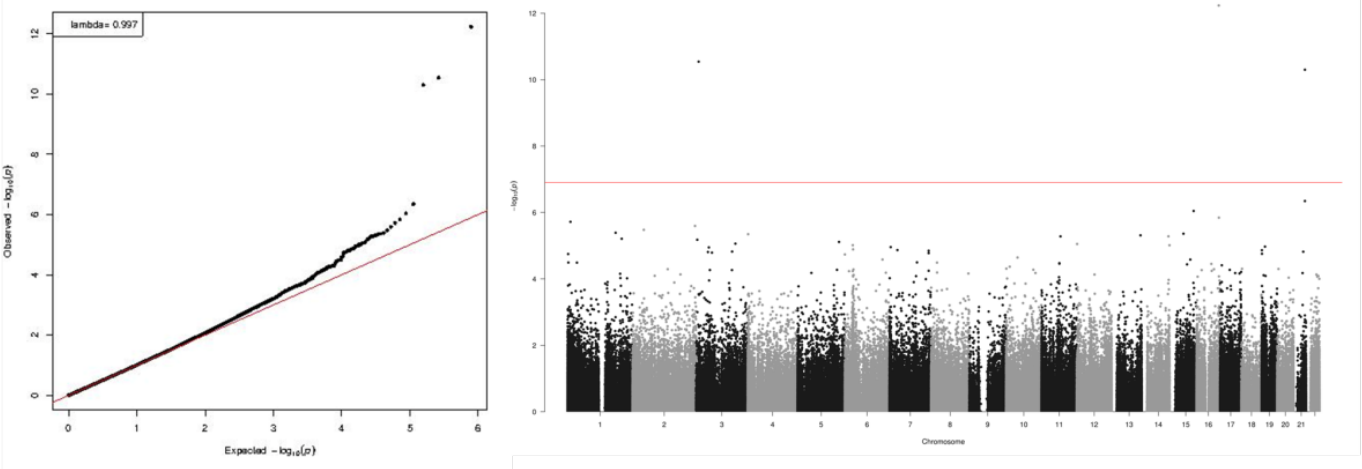
B



C

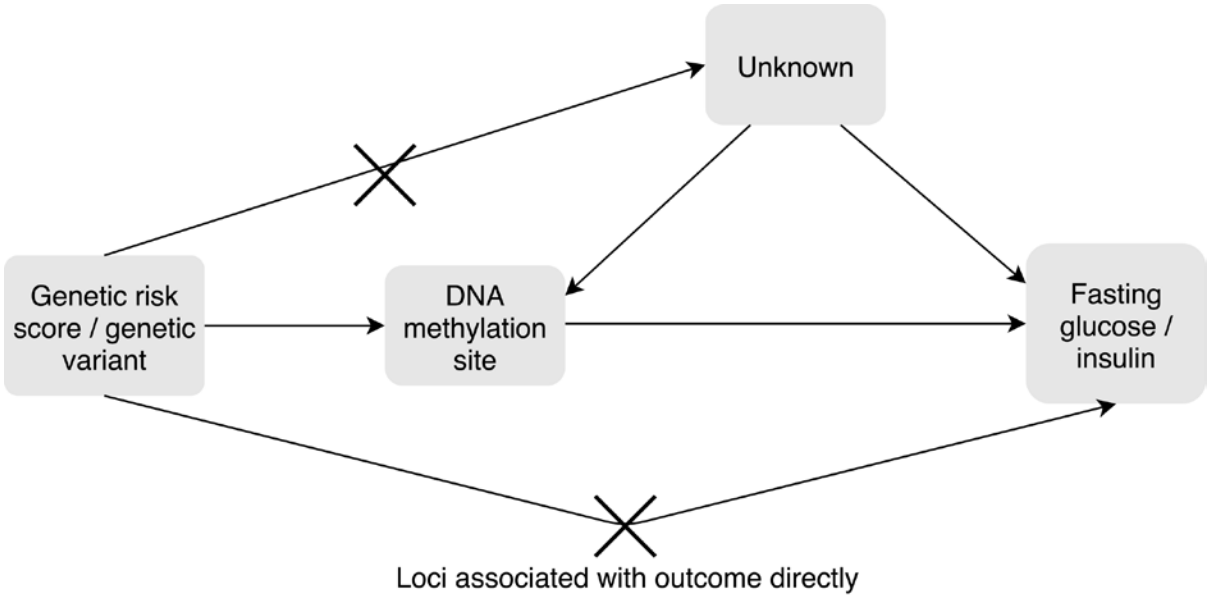


D



Supplementary Figure 1 QQ plots and Manhattan plots of the epigenome-wide association study (EWAS) results

A: EWAS results of fasting glucose in the baseline model; B: EWAS results of fasting glucose in the BMI-adjusted model; C: EWAS results of fasting insulin in the baseline model; D: D EWAS results of fasting insulin in the BMI-adjusted model.



Supplementary Figure 2 Overview of the general Mendelian Randomization process

Disclaimer

The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

References

1 Meigs, J. B., Nathan, D. M., D'Agostino, R. B., Sr., Wilson, P. W. & Framingham Offspring, S. Fasting and postchallenge glycemia and cardiovascular disease risk: the Framingham Offspring Study. *Diabetes Care* **25**, 1845-1850 (2002).

2 Coutinho, M., Gerstein, H. C., Wang, Y. & Yusuf, S. The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. *Diabetes Care* **22**, 233-240 (1999).

3 Matthews, D. R. *et al.* Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412-419 (1985).

4 Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* **42**, 105-116, doi:10.1038/ng.520 (2010).

5 Manning, A. K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet* **44**, 659-669, doi:10.1038/ng.2274 (2012).

6 Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* **44**, 991-1005, doi:10.1038/ng.2385 (2012).

7 Scott, R. A. *et al.* An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* **66**, 2888-2902, doi:db16-1253 [pii]
10.2337/db16-1253 (2017).

8 Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* **44**, 981-990, doi:10.1038/ng.2383 (2012).

9 Replication, D. I. G. *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* **46**, 234-244, doi:10.1038/ng.2897 (2014).

10 Voight, B. F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* **42**, 579-589, doi:10.1038/ng.609 (2010).

11 Ali, O. Genetics of type 2 diabetes. *World J Diabetes* **4**, 114-123, doi:10.4239/wjd.v4.i4.114 (2013).
12 Hidalgo, B. *et al.* Epigenome-wide association study of fasting measures of glucose, insulin, and HOMA-IR in the Genetics of Lipid Lowering Drugs and Diet Network study. *Diabetes* **63**, 801-807 (2014).
13 Wahl, S. *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81-86 (2017).
14 Kahn, S. E., Hull, R. L. & Utzschneider, K. M. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* **444**, 840-846, doi:nature05482 [pii]
10.1038/nature05482 (2006).
15 Xi, B. *et al.* Associations of genetic variants in/near body mass index-associated genes with type 2 diabetes: a systematic meta-analysis. *Clin Endocrinol (Oxf)* **81**, 702-710, doi:10.1111/cen.12428 (2014).
16 Ehrlich, M. & Lacey, M. DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. *Epigenomics* **5**, 553-568, doi:10.2217/epi.13.43 (2013).
17 Al Muftah, W. A. *et al.* Epigenetic associations of type 2 diabetes and BMI in an Arab population. *Clin Epigenetics* **8**, 13 (2016).
18 Aslibekyan, S. *et al.* Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity (Silver Spring)* **23**, 1493-1501 (2015).
19 Chambers, J. C. *et al.* Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol* **3**, 526-534 (2015).
20 Demerath, E. W. *et al.* Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet* **24**, 4464-4479 (2015).
21 Ding, J. *et al.* Alterations of a Cellular Cholesterol Metabolism Network Are a Molecular Feature of Obesity-Related Type 2 Diabetes and Cardiovascular Disease. *Diabetes* **64**, 3464-3474 (2015).
22 Kriebel, J. *et al.* Association between DNA Methylation in Whole Blood and Measures of Glucose Metabolism: KORA F4 Study. *PLoS One* **11**, e0152314 (2016).
23 Kulkarni, H. *et al.* Novel epigenetic determinants of type 2 diabetes in Mexican-American families. *Hum Mol Genet* **24**, 5330-5344 (2015).
24 Wang, B. *et al.* Methylation loci associated with body mass index, waist circumference, and waist-to-hip ratio in Chinese adults: an epigenome-wide analysis. *Lancet* **388 Suppl 1**, S21 (2016).
25 Wilson, L. E., Harlid, S., Xu, Z., Sandler, D. P. & Taylor, J. A. An epigenome-wide study of body mass index and DNA methylation in blood using participants from the Sister Study cohort. *Int J Obes (Lond)* **41**, 194-199 (2017).
26 Bonder, M. J. *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* **49**, 131-138 (2017).

27 Mahajan, A. *et al.* Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet* **50**, 559-571 (2018).
28 Barbeira, A. *et al.* Integrating tissue specific mechanisms into GWAS summary results. *bioRxiv*, 045260 (2016).
29 Barbeira, A. *et al.* MetaXcan: Summary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results. *bioRxiv*, 045260 (2016).
30 Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet* **8**, e1002607 (2012).
31 Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* **47**, 1091-1098, doi:ng.3367 [pii]
10.1038/ng.3367 (2015).
32 Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585 (2013).
33 Aguet, F. *et al.* Local genetic effects on gene expression across 44 human tissues. *BiorXiv*, 074450 (2016).
34 Grarup, N. *et al.* Association of variants in the sterol regulatory element-binding factor 1 (SREBF1) gene with type 2 diabetes, glycemia, and insulin resistance: a study of 15,734 Danish subjects. *Diabetes* **57**, 1136-1142, doi:db07-1534 [pii]
10.2337/db07-1534 (2008).
35 Armstrong, D. L. *et al.* GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. *Genes Immun* **15**, 347-354, doi:gene201423 [pii]
10.1038/gene.2014.23 (2014).
36 Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119-124, doi:nature11582 [pii]
10.1038/nature11582 (2012).
37 Reiner, A. P. *et al.* Genome-wide association and population genetic analysis of C-reactive protein in African American and Hispanic American women. *Am J Hum Genet* **91**, 502-512, doi:S0002-9297(12)00409-0 [pii]
10.1016/j.ajhg.2012.07.023 (2012).
38 Wu, T., Dorn, J. P., Donahue, R. P., Sempos, C. T. & Trevisan, M. Associations of serum C-reactive protein with fasting insulin, glucose, and glycosylated hemoglobin: the Third National Health and Nutrition Examination Survey, 1988-1994. *Am J Epidemiol* **155**, 65-71 (2002).
39 van der Harst, P. *et al.* 52 Genetic Loci Influencing Myocardial Mass. *J Am Coll Cardiol* **68**, 1435-1448, doi:S0735-1097(16)34664-2 [pii]
10.1016/j.jacc.2016.07.729 (2016).
40 Conti, D. V. *et al.* Two Novel Susceptibility Loci for Prostate Cancer in Men of African Ancestry. *J Natl Cancer Inst* **109**, doi:3858844 [pii]

10.1093/jnci/djx084 (2017).
41 Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427, doi:nature13595 [pii]
10.1038/nature13595 (2014).
42 Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-365, doi:10.1038/366362a0 (1993).
43 Ferre, P. & Foufelle, F. Hepatic steatosis: a role for de novo lipogenesis and the transcription factor SREBP-1c. *Diabetes Obes Metab* **12 Suppl 2**, 83-92, doi:10.1111/j.1463-1326.2010.01275.x (2010).
44 Huang, Y. T. *et al.* Epigenome-wide profiling of DNA methylation in paired samples of adipose tissue and blood. *Epigenetics* **11**, 227-236, doi:10.1080/15592294.2016.1146853 (2016).
45 Ikram, M. A. *et al.* The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol* **32**, 807-850, doi:10.1007/s10654-017-0321-4
10.1007/s10654-017-0321-4 [pii] (2017).
46 Boomsma, D. I. *et al.* Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* **9**, 849-857, doi:10.1375/183242706779462426
S1832427400007192 [pii] (2006).
47 van Dongen, J., Willemsen, G., Chen, W. M., de Geus, E. J. & Boomsma, D. I. Heritability of metabolic syndrome traits in a large population-based sample. *J Lipid Res* **54**, 2914-2923, doi:jlr.P041673 [pii]
10.1194/jlr.P041673 (2013).
48 Moayyeri, A., Hammond, C. J., Hart, D. J. & Spector, T. D. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet* **16**, 144-149 (2013).
49 Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86, doi:1471-2105-13-86 [pii]
10.1186/1471-2105-13-86 (2012).
50 Chen, Y. A. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203-209, doi:23470 [pii]
10.4161/epi.23470 (2013).
51 Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* **36**, 1-48 (2010).
52 Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
53 Dayem Ullah, A. Z., Lemoine, N. R. & Chelala, C. SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic Acids Res* **40**, W65-70, doi:gks364 [pii]

10.1093/nar/gks364 (2012).
54 Chelala, C., Khan, A. & Lemoine, N. R. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics* **25**, 655-661, doi:btn653 [pii]
10.1093/bioinformatics/btn653 (2009).
55 Liu, J. *et al.* A Mendelian Randomization Study of Metabolite Profiles, Fasting Glucose, and Type 2 Diabetes. *Diabetes* **66**, 2915-2926, doi:db17-0199 [pii]
10.2337/db17-0199 (2017).
56 International Consortium for Blood Pressure Genome-Wide Association, S. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103-109, doi:nature10405 [pii]
10.1038/nature10405 (2011).
57 Soranzo, N. *et al.* Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* **59**, 3229-3239, doi:10.2337/db10-0502 (2010).
58 Wheeler, E. *et al.* Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med* **14**, e1002383, doi:10.1371/journal.pmed.1002383 PMEDICINE-D-17-00599 [pii] (2017).