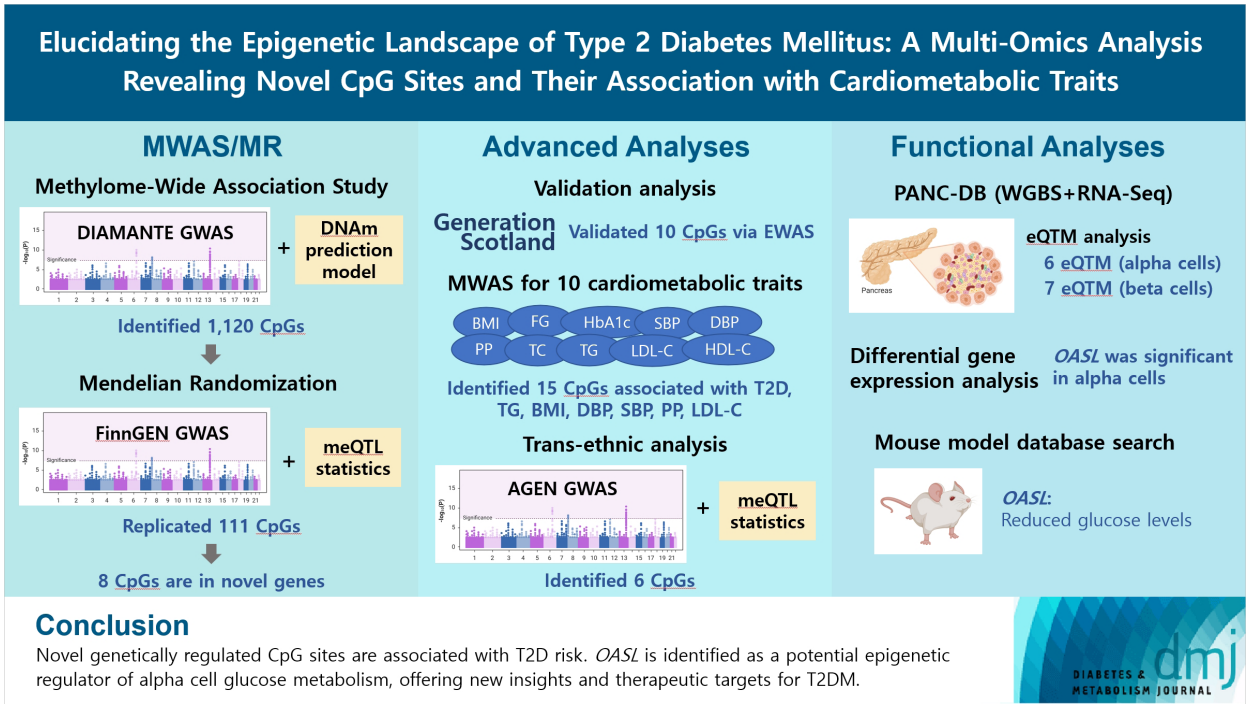
dmj
DIABETES & METABOLISM JOURNAL

Elucidating the Epigenetic Landscape of Type 2 Diabetes Mellitus: A Multi-Omics Analysis Revealing Novel CpG Sites and Their Association with Cardiometabolic Traits

Ren-Hua Chung, Chun-Chao Wang, Djeane Debora Onthoni, Ben-Yang Liao, Tzu-Sheng Hsu, Eden R. Martin, Chao A. Hsiung, Wayne Huey-Herng Sheu, Hung-Yi Chiou

Diabetes Metab J 2026;50:153-164 | *https://doi.org/10.4093/dmj.2025.0041*



Highlights

- **We identified 111 CpGs linked to T2DM in Europeans, including 8 novel CpG sites.**
- **Six CpGs showed trans-ethnic effects, replicated in East Asian populations.**
- **OASL gene expression was altered in T2DM alpha cells, linked to CpG methylation.**
- **Many CpGs overlapped with cardiometabolic traits like TG and blood pressure.**
- **Mouse data supported OASL's role in regulating glucose metabolism in T2DM.**

How to cite this article:
Chung RH, Wang CC, Onthoni DD, Liao BY, Hsu TS, Martin ER, et al. Elucidating the Epigenetic Landscape of Type 2 Diabetes Mellitus: A Multi-Omics Analysis Revealing Novel CpG Sites and Their Association with Cardiometabolic Traits. Diabetes Metab J 2026;50:153-164. https://doi.org/10.4093/dmj.2025.0041

Original Article
Genetics

Diabetes Metab J 2026;50:153-164
https://doi.org/10.4093/dmj.2025.0041
pISSN 2233-6079 · eISSN 2233-6087

dmj
DIABETES & METABOLISM JOURNAL

Elucidating the Epigenetic Landscape of Type 2 Diabetes Mellitus: A Multi-Omics Analysis Revealing Novel CpG Sites and Their Association with Cardiometabolic Traits

Ren-Hua Chung[1], Chun-Chao Wang[2], Djeane Debora Onthoni[1], Ben-Yang Liao[1], Tzu-Sheng Hsu[3], Eden R. Martin[4,5], Chao A. Hsiung[1], Wayne Huey-Herng Sheu[6], Hung-Yi Chiou[1,7]

[1]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan
[2]Institute of Molecular Medicine & Department of Medical Science, College of Life Sciences and Medicine, National Tsing Hua University, Hsinchu, Taiwan
[3]Institute of Molecular and Cellular Biology, College of Life Sciences and Medicine, National Tsing Hua University, Hsinchu, Taiwan
[4]John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA
[5]John T. MacDonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA
[6]Institute of Molecular and Genomic Medicine, National Health Research Institutes, Zhunan, Taiwan
[7]School of Public Health, College of Public Health, Taipei Medical University, Taipei, Taiwan

Background: Type 2 diabetes mellitus (T2DM) is a complex, multifactorial disease with a significant global burden. Although genome-wide association studies (GWAS) have identified many T2DM-associated variants, most lie in non-coding regions, making it difficult to interpret their functional roles.
Methods: We aimed to identify genetically regulated Cytosine–phosphate–Guanine (CpG) sites associated with T2DM by conducting a methylome-wide association study (MWAS), followed by Mendelian randomization (MR) and functional validation using human pancreatic cells and mouse models. MWAS was performed using summary statistics from large-scale GWAS and a DNA methylation (DNAm) prediction model to test associations between genetically predicted DNAm and T2DM.
Results: We identified 111 CpG sites significantly associated with T2DM in Europeans, including 8 novel sites near genes not previously linked to T2DM. These findings were replicated in independent datasets. Many CpGs also showed associations with cardiometabolic traits, highlighting shared epigenetic mechanisms. Trans-ethnic MR analysis confirmed consistent effects for six CpGs in East Asians. Functional analysis revealed that several CpGs regulate gene expression in human pancreatic α- and β-cells. Among them, 2′-5′-oligoadenylate synthetase like (*OASL*) expression, regulated by a significant CpG, was differentially expressed in α-cells of T2DM cases compared to controls. Supporting evidence from mouse models suggests a role for *OASL* in glucose regulation.
Conclusion: Our study identifies novel genetically regulated CpG sites associated with T2DM risk and highlights *OASL* as a potential epigenetic regulator of glucose metabolism in α-cells. These findings provide mechanistic insights into the epigenetic architecture of T2DM and suggest potential targets for cross-ethnic biomarker development and therapeutic intervention.

Keywords: Diabetes mellitus, type 2; DNA methylation; Epigenomics; Genome-wide association study; Mendelian randomization analysis

Corresponding author: Ren-Hua Chung https://orcid.org/0000-0002-9835-6333
Institute of Population Health Sciences, National Health Research Institutes,
No 35, Keyan Road, Zhunan, Miaoli 350, Taiwan
E-mail: rchung@nhri.edu.tw

Received: Jan. 13, 2025; Accepted: May 19, 2025

INTRODUCTION

Type 2 Diabetes (T2DM) is a complex multifactorial disease with significant global health impact. Genome-wide association studies (GWAS) have identified numerous genetic variants associated with T2DM [1], but understanding the functional mechanisms by which these variants contribute to the disease remains challenging. This difficulty arises from the polygenic nature of T2DM, the complexity of distinguishing causal variants from those in linkage disequilibrium (LD), and the fact that most GWAS hits are located in non-coding regions of the genome [2].

Integrated approaches combining GWAS data with functional genomics have been employed to address these challenges [3]. One such approach is transcriptome-wide association studies (TWAS), which leverage expression quantitative trait loci (eQTL) to impute gene expression levels and link them to phenotypic traits [4,5]. TWAS has successfully uncovered novel gene-disease associations in complex diseases, including T2DM, providing insights into disease etiology and highlighting potential therapeutic targets [6,7].

The principles underlying TWAS are effectively applicable to the study of DNA methylation (DNAm), a pivotal mechanism for regulation of gene expression. DNAm involves the addition of a methyl group to cytosine residues within Cytosine–phosphate–Guanine (CpG) sites, influencing gene activity by altering the accessibility of transcriptional machinery [8]. Abnormal DNAm levels are associated with various complex diseases, including T2DM, as identified through epigenome-wide association studies (EWAS) [9]. Importantly, DNAm is not just an epigenetic phenomenon but also a heritable one. Studies have demonstrated that CpG methylation possesses a significant heritable component, with heritability estimates ranging from 16% to 20% [10]. Consequently, many methylation quantitative trait loci (meQTL), which are genetic variants affecting the DNAm levels, have been identified [11]. The combination of DNAm's heritability and the availability of meQTL data paves the way for methylome-wide association studies (MWAS). MWAS tests the association between imputed DNAm levels—based on meQTL from GWAS—and diseases, providing a powerful approach to uncover genetically regulated DNAm associations with diseases and offering insights into disease mechanisms that may not be detectable through traditional GWAS or TWAS methods.

Despite its potential, the application of MWAS in T2DM remains unexplored, and studies focusing on non-European populations are notably absent. Fryett et al. [10] pioneered the construction of a methylation prediction model and identified over 700 significant CpGs in primary biliary cholangitis. However, similar studies in T2DM and diverse populations are lacking. This gap highlights the need for MWAS to uncover genetically regulated DNAm associations with T2DM, particularly in underrepresented populations.

To address this gap, we conducted a MWAS to identify CpGs where DNAm levels regulated by meQTL are associated with T2DM risk in Europeans. We replicated significant findings using Mendelian randomization (MR) in independent datasets and validated them with measured methylation levels. Additionally, we evaluated the associations of these CpGs with other cardiometabolic traits and assessed their trans-ethnic effects in East Asians by leveraging large-scale genomic data from the Taiwan Biobank (TWB) [12] and the Asian Genetic Epidemiology Network (AGEN) consortium [13]. The inclusion of these datasets strengthens the generalizability of our findings by extending the analysis beyond European populations, addressing the scarcity of MWAS in East Asians. Recognizing that both pancreatic β-cells, which produce insulin, and α-cells, which secrete glucagon, play significant roles in T2DM pathophysiology [14], we integrated methylation and gene expression data from human pancreatic α- and β-cells to assess the associations of gene expression regulated by CpGs identified in our study with T2DM. Dysregulated glucagon secretion from α-cells contributes to hyperglycemia by promoting hepatic glucose production [15], and DNAm patterns in α-cells may influence gene expression related to glucagon secretion and glucose metabolism [16]. Therefore, investigating epigenetic modifications in both α- and β-cells can provide a more comprehensive understanding of T2DM mechanisms. By evaluating functional implications using databases of mouse models, we aimed to elucidate the biological mechanisms underlying these associations. Our study provides novel insights into the epigenetic regulation of T2DM and highlights potential therapeutic targets.

METHODS

Overview of datasets and models

A comprehensive summary of the datasets and models used in this study is provided in Supplementary Table 1, including details on sample size, ancestry, data type, and references. Our

analyses primarily relied on large-scale GWAS summary statistics for T2DM, including data from the Diabetes Meta-analysis of Trans-ethnic Association Studies (DIAMANTE) consortium (80,154 T2DM cases and 853,816 controls of European ancestry) [17] and the FinnGen study (38,657 cases and 310,131 controls) [18]. For the MWAS, we employed a DNAm prediction model trained using genotype and methylation data from the Understanding Society cohort (n=1,120) [19]. meQTL information was obtained from the MeQTL EPIC database (n=2,358) [11], which provided single-nucleotide polymorphism (SNP)-CpG associations profiled using the Illumina EPIC array (San Diego, CA, USA). We performed MWAS analyses across ten cardiometabolic traits using trait-specific GWAS summary statistics from European populations obtained through the GWAS catalog. To assess trans-ethnic generalizability, we incorporated East Asian datasets, including meQTL summary statistics from the TWB (n=2,150) [12] and T2DM GWAS summary statistics from the AGEN consortium (77,418 cases and 356,122 controls) [13]. Functional validation was conducted using methylation and gene expression data from human pancreatic α- and β-cells in Data portal of The Human Pancreas Analysis Program (PANC-DB) (https://hpap.pmacs.upenn.edu) [20], and further biological insights were derived from the Mouse Genome Informatics (MGI) database.

MWAS for type 2 diabetes mellitus

We performed a MWAS to identify CpG sites associated with T2DM using S-PrediXcan from the MetaXcan software [21], a tool designed for TWAS that leverages GWAS summary statistics to infer associations between imputed molecular traits (e.g., methylation levels) and phenotypes. This method does not require individual-level genotype or methylation data, making it particularly useful for large-scale studies.

S-PrediXcan estimates the association between genetically predicted methylation levels and T2DM by integrating: (1) a methylation prediction model trained on an independent reference dataset, which imputes methylation levels using nearby meQTLs, and (2) GWAS summary statistics, which provide association signals between these meQTLs and T2DM. This approach effectively links methylation regulation to disease risk while accounting for LD among genetic variants.

The methylation prediction model used in this study was developed by Fryett et al. [10], based on data from the Understanding Society cohort [19], which includes genotype and methylation data from 1,120 individuals. The model applies penalized regression techniques, including least absolute shrinkage and selection operator (LASSO), ridge regression, and elastic net, to predict methylation levels (i.e., β values, which represent the proportion of methylated alleles at a particular CpG site). These methods use genotypic data at nearby SNPs within a specific window as predictors for methylation levels at CpG sites. The models were optimized using cross-validation to identify the optimal window sizes and penalization parameters to maximize prediction accuracy. Prediction accuracy, measured by the correlation between observed and predicted methylation levels, was validated using an independent dataset from the Accessible Resource for Integrated Epigenomics Studies (ARIES) [22], which consists of 841 samples. In total, the model predicts methylation levels at over 179,000 CpG sites, achieving a prediction accuracy greater than 0.1 in both the training and validation datasets. Most of the CpGs selected by this threshold achieved the upper bound of heritability [10], ensuring the reliability of the predicted methylation levels for downstream association testing.

We used GWAS summary statistics from the DIAMANTE consortium, comprising 80,154 T2DM cases and 853,816 controls of European ancestry [17]. The dataset was obtained from the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) website (https://diagram-consortium.org/downloads.html) under the file name 'Ancestry-specific GWAS meta-analysis summary statistics: European.' Using S-PrediXcan, we tested the association of methylation levels at 179,523 CpG sites with T2DM. CpG sites with association P values below 2.78×10^{-7} were considered significant after Bonferroni correction for multiple testing. Further details on the S-PrediXcan algorithm and its application in MWAS are provided in the Supplementary Methods.

Replication of MWAS results using MR analysis

To strengthen the findings from our MWAS and establish a causal relationship between DNAm and T2DM, we used MR for replication. MR is particularly well-suited for addressing limitations in MWAS by helping to infer causality rather than mere association [23]. In the context of our study, MR allows us to test whether the CpGs identified as significant in MWAS have a causal effect on T2DM, independent of confounding factors such as pleiotropy (where a genetic variant affects multiple traits via different biological pathways).

In this study, we conducted a two-sample MR analysis using the inverse-variance weighted (IVW) [24] algorithm, the most

widely used method in MR, alongside the MR-Egger [25] intercept test to evaluate the potential effect of pleiotropy. Both IVW and MR-Egger tests were conducted using the R package MendelianRandomization v0.10.1 (R Foundation for Statistical Computing, Vienna, Austria) [26]. The F-statistic [27] and Cochran's Q statistic, derived from the IVW estimate, and its associated *P* value were also calculated using MendelianRandomization v0.10.1 to assess weak instrument bias and heterogeneity.

For the MR analysis, we selected CpGs that were significant in the MWAS and used meQTL as instrumental variables. The association signals for these meQTL with CpGs were obtained from the MeQTL EPIC database [11], which contains data on 724,499 CpGs profiled on the Illumina Infinium MethylationEPIC array in 2,358 blood samples from three UK cohorts. We chose clumped meQTL with low LD for independence and applied a false discovery rate <0.05 to exclude weak instrumental variables, ensuring robust MR analysis.

Association statistics of the instrumental variables with T2DM were obtained from the FinnGen study [18], another large-scale GWAS for T2DM in Europeans, consisting of 38,657 T2DM cases and 310,131 controls. Importantly, the meQTL and T2DM GWAS datasets used in the MR analysis were independent from those used in the initial MWAS, ensuring unbiased replication of the signals. In adhering to the key assumptions of two-sample MR, we ensured that the summary statistics for the Exposure-SNP (i.e., meQTL information) and Outcome-SNP (i.e., T2DM associations) were derived from independent but related populations. This approach aligns with the principle of using non-overlapping datasets for the exposure and outcome in two-sample MR to avoid bias.

Validation of the CpGs using measured DNAm levels

To validate the signals identified from the MWAS and MR analyses, we compared our findings against external studies based on measured DNAm levels. For this purpose, we used the EWAS Catalog [9], which compiles association statistics for CpGs and traits from published EWAS. We downloaded the full summary statistics from the Generation Scotland (GS) cohort [28], calculated based on 757 incident T2DM cases and 16,992 controls—the largest T2DM EWAS dataset available in the catalog.

MWAS for cardiometabolic traits

To further explore the broader implications of the CpGs identified in our study, we extended our analysis to investigate their effects on cardiometabolic traits that are closely related to T2DM. These traits included body mass index (BMI), fasting glucose, glycosylated hemoglobin, systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse pressure (PP), total cholesterol, triglycerides (TG), low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C). We accessed GWAS summary statistics for these traits in European populations from the GWAS Catalog [29-33]. Using S-PrediXcan, we then conducted MWAS analyses for each trait, employing Fryett's DNAm prediction model alongside the respective trait's GWAS summary statistics to assess whether the CpGs identified for T2DM also play a role in these related cardiometabolic traits.

Trans-ethnic analysis of the significant CpGs

To determine the broader applicability and generalizability of our findings beyond the European population, we conducted a trans-ethnic analysis to assess whether the CpGs identified in our study exhibit consistent effects on T2DM among East Asians. This analysis aimed to explore whether the relationships between these CpGs and T2DM hold across different ethnicities, which is crucial for validating the global relevance of these CpGs. We also used a two-sample MR analysis to investigate the causal influence of these CpGs on T2DM. The meQTL summary statistics were derived from 2,150 samples profiled based on the Illumina Infinium MethylationEPIC BeadChip, with GWAS data from the TWB [12]. Supplementary Methods provide in-depth details regarding the QC procedures for the TWB samples and the methodology for computing the meQTL summary statistics. GWAS summary statistics for T2DM were obtained from the AGEN consortium, based on a cohort of 77,418 T2DM cases and 356,122 controls [13].

Expression quantitative trait methylation and differential gene expression analyses

To investigate whether the significant CpGs function as expression quantitative trait methylation (eQTM) loci in human pancreatic α- and β-cells, we utilized whole-genome bisulfite sequencing (WGBS) and RNA sequencing (RNA-Seq) data from PANC-DB [20]. Given the essential role of α-cells in glucagon secretion and β-cells in insulin secretion, examining DNAm patterns in both cell types is crucial, as their dysfunction contributes to T2DM pathophysiology [34]. The processing details for the WGBS and RNA-Seq data are elaborated in the Supplementary Methods. Our analysis was confined to Eu-

ropean-descent samples without a diabetes history, comprising 17 α-cell samples and 10 β-cell samples with both WGBS and RNA-Seq data.

For eQTM analysis, we modeled the transcript per million (TPM) for each gene using a linear regression model, accounting for age and sex, with R version 4.3.2. The residuals from this model were then correlated with methylation levels using Kendall's τ coefficient, with significance determined by a non-parametric τ test. Given that CpGs often regulate gene expression from positions upstream, our focus was on CpGs located within 100 kb upstream of genes. To maximize the identification of potential genes regulated by CpGs, we did not apply multiple testing correction at this stage. Instead, genes with correlation $P<0.05$ were selected for further analysis, and multiple testing correction was performed in the subsequent differential gene expression analysis to ensure robust findings.

To explore the expression dynamics of genes potentially regulated by the CpGs identified in the eQTM analysis, we conducted a differential gene expression analysis between T2DM cases and controls without a diabetes history. This analysis utilized RNA-Seq data from PANC-DB, focusing on samples from pancreatic α- and β-cells. Specifically, the dataset included six T2DM patients and 26 controls for α-cells, and six T2DM patients and 22 controls for β-cells. We first adjusted the TPM for each gene by age and sex using a linear regression model, then derived residuals for further analysis. These residuals were assessed for their association with T2DM status employing a permutation test. Detailed methodology of the permutation test is elaborated in the Supplementary Methods. Genes that exhibited significant associations with T2DM status, as determined by permutation *P* values and further adjusted for multiple testing using Bonferroni correction, were earmarked for subsequent functional analysis, aiming to uncover the biological implications of these expression differences in the context of T2DM.

Functional analysis using existing mouse model data

To delve deeper into the functional significance of the genes identified in the differential gene expression analysis, we searched the MGI database. This comprehensive resource integrates data from the Mouse Genome Database (MGD) [35], Mouse Gene Expression Database (GXD) [36], and Mouse Tumor Biology Database (MMHCdb) [37]. Our investigation focused on examining the phenotypic outcomes of mutational changes in these genes, aiming to infer potential implications for diabetes pathogenesis and progression. This approach allows us to extrapolate the biological impact of these genes based on phenotypic alterations observed in mouse models, providing valuable insights into their roles in disease mechanisms.

Ethnics approval and consent to participate

Written informed consent for participants in the TWB was obtained from all participants. Research in this study was approved by the Institutional Review Board of the National Health Research Institutes in Taiwan (reference number: EC1091202-E).

RESULTS

Our analysis workflow is illustrated in Fig. 1, and the resources utilized throughout our study are detailed in Supplementary Table 1. The MWAS and MR analyses primarily leveraged GWAS summary statistics from datasets comprising hundreds of thousands of samples.

Identification and replication of CpGs

The MWAS analysis identified 1,120 significant CpGs. For MR analysis, we required a minimum of three instrumental meQTL per CpG to enable pleiotropy tests, resulting in 495 out of the 1,120 CpGs being eligible for further analysis. Of these, 111 CpGs were successfully replicated, showing significant MR $P<1.01\times10^{-4}$—considering the 495 tests conducted—without evidence of directional pleiotropy (MR-Egger $P>0.05$) or heterogeneity among instrumental variables (Cochran's Q statistic $P>0.05$). Moreover, all 111 CpGs had F-statistics >10, indicating that the instrumental variables were sufficiently strong and unlikely to introduce weak instrument bias.

Novelty and genomic context of replicated CpGs

Among the 111 replicated CpGs, eight CpGs are located within or proximal to genes not previously highlighted as T2DM candidate genes in the GWAS catalog or the Type 2 Diabetes Knowledge Portal [38], where candidate genes are defined as those with a human genetic evidence (HuGE) score ≥10 [39]. Detailed MWAS and MR statistics for these eight novel CpGs are presented in Table 1, whereas comprehensive details regarding all 111 CpGs can be found in Supplementary Table 2. Across MWAS and MR analyses, the direction of effect generally concurs, where positive effects indicated that higher genetically predicted DNAm levels increased T2DM risk.

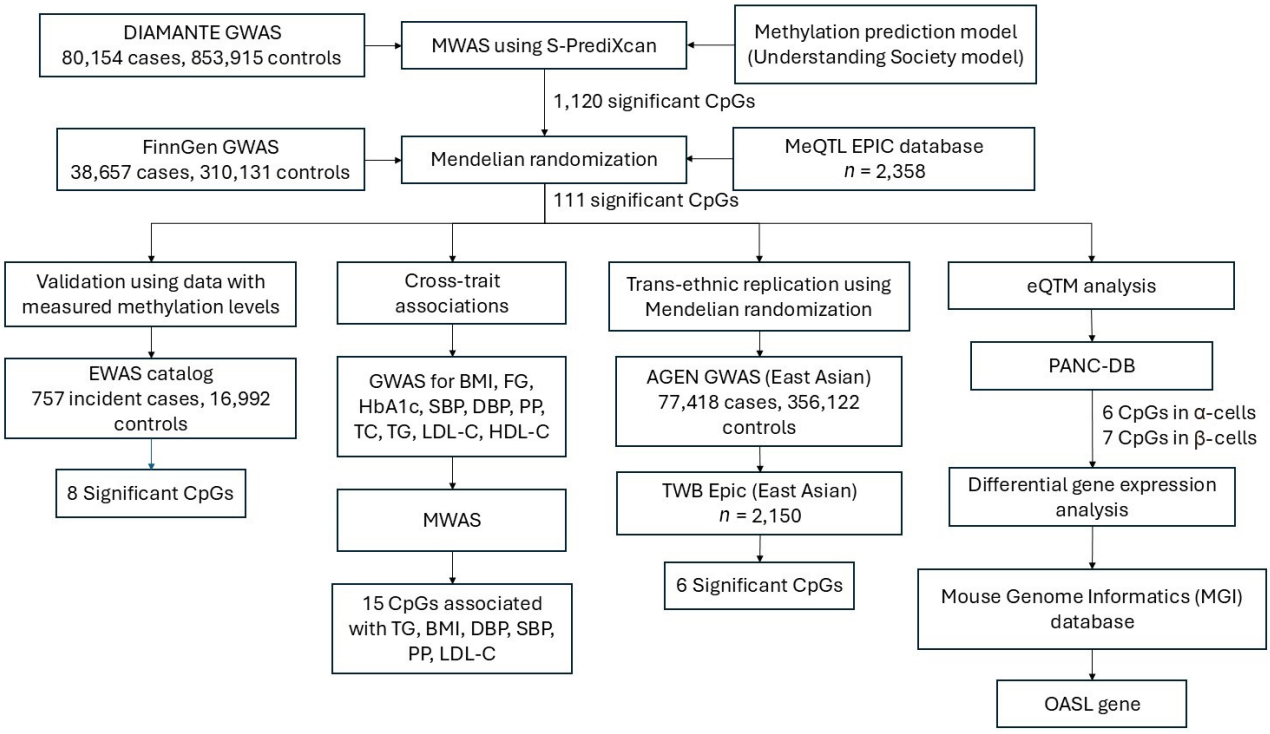


Fig. 1. Study overview and key findings. Methylome-wide association study (MWAS): Using S-PrediXcan and Diabetes Meta-analysis of Trans-ethnic Association Studies (DIAMANTE) genome-wide association study (GWAS) summary statistics (80,154 cases and 853,915 controls), 1,120 cytosine-guanine dinucleotides (CpGs) associated with type 2 diabetes mellitus (T2DM) were identified in Europeans. Mendelian randomization (MR): Among them, 111 CpGs were supported by MR analysis using independent FinnGen GWAS data (38,657 cases and 310,131 controls). Validation with measured methylation data: Of the 111 CpGs, eight were validated using measured DNA methylation data from the Generation Scotland (GS) cohort from the epigenome-wide association study (EWAS) catalog (757 cases and 16,992 controls). Cross-trait associations: MWAS analysis of the 111 CpGs with 10 cardiometabolic traits identified 15 CpGs showing significant associations with T2DM, triglycerides (TG), body mass index (BMI), diastolic blood pressure (DBP), systolic blood pressure (SBP), pulse pressure (PP), and low-density lipoprotein cholesterol (LDL-C). Trans-Ethnic Replication: Using Asian Genetic Epidemiology Network (AGEN) East Asian GWAS data (77,418 cases and 356,122 controls) and methylation data from the Taiwan Biobank Illumina Infinium MethylationEPIC BeadChip (Taiwan Biobank [TWB] EPIC) (n=2,150), six CpGs demonstrated trans-ethnic effects. Functional analyses: Expression quantitative trait methylation (eQTM) analysis using Data portal of The Human Pancreas Analysis Program (PANC-DB) identified six CpGs in α-cells and seven in β-cells. Differential gene expression analysis revealed that 2′-5′-oligoadenylate synthetase like (*OASL*) was significantly regulated in α-cells. Functional relevance of *OASL* in glucose homeostasis was supported by data from the Mouse Genome Informatics (MGI) database. MeQTL, methylation quantitative trait loci; FG, fasting glucose; HbA1c, glycosylated hemoglobin; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol.

Genetic variants associated with replicated CpGs

Supplementary Table 3 outlines the lead SNPs—those with the smallest GWAS *P* values—associated with each of the 111 CpGs, identified either in the MWAS as prediction variables or in the MR analysis as instrumental variables. These lead SNPs generally exhibited significant GWAS $P<5\times10^{-8}$. The median proximity of these SNPs to the corresponding CpGs was approximately 27 kb in the MWAS context and about 51 kb in the MR analysis.

Validation of CpGs using measured DNAm data

Out of the 111 identified CpG sites, eight were successfully validated through measured DNAm levels from the GS study, with $P<\frac{0.05}{111}=4.50\times10^{-4}$. Their summary statistics are shown in Table 2. The validation of these CpGs, despite the relatively modest cohort size of 757 T2DM cases, suggests that the genetically regulated aspects of methylation at these eight sites have a substantial impact.

Table 1. Significant CpGs in novel genes for type 2 diabetes mellitus

CpG	Chromosome	Position	Gene name	MWAS effect size	MWAS *P* value	SNPs in model[a]	MR estimate (SD)	MR *P* value	Pleiotropy *P* value[b]	Heterogeneity *P* value[c]	SNPs used in MR[d]
cg27359689	5	75034394	*LOC441087-SV2C*	1.778	2.90E-10	11	0.039 (0.008)	2.07E-06	6.31E-01	2.25E-01	3
cg16837680	8	8084974	*FAM85B*	3.202	2.12E-07	12	0.069 (0.016)	8.90E-06	2.45E-01	8.00E-01	5
cg01236304	8	10614369	*SOX7-AS1*	–1.159	4.18E-09	23	–0.035 (0.005)	1.10E-12	5.30E-01	1.43E-01	33
cg10813958	8	11247920	*FAM167A-AS1*	–2.390	1.91E-09	20	–0.050 (0.009)	8.15E-09	6.29E-01	3.38E-01	18
cg02034205	8	11382291	*BLK*	4.248	1.58E-07	13	0.034 (0.008)	1.41E-05	8.96E-01	1.46E-01	21
cg05622566	11	17058241	*PLEKHA7-OR7E14P*	4.885	1.20E-07	19	0.101 (0.025)	7.14E-05	8.73E-01	2.57E-01	3
cg21805149	14	91871665	*CCDC88C*	–4.877	4.33E-09	44	–0.085 (0.019)	1.02E-05	4.98E-01	5.44E-01	3
cg08228953	19	7979674	*TGFBR3L*	–3.003	7.36E-11	39	–0.078 (0.012)	3.33E-11	6.73E-01	1.70E-01	3

CpG, cytosine-guanine dinucleotide; MWAS, methylome-wide association study; SNP, single-nucleotide polymorphism; MR, Mendelian randomization; SD, standard deviation.
[a]The number of SNPs used in the DNA methylation prediction model for the MWAS, [b]The *P* value of the intercept test from MR-Egger, [c]The *P* value of the Cochran's Q statistic, [d]The number of instrumental variables used in the MR analysis.

Table 2. CpGs that were validated using samples with measured DNA methylation levels

CpG	Chromosome	Position	Gene name	β	SE	*P* value
cg06697744	1	39680237	*MACF1*	0.053	0.014	7.96E-05
cg20885688	5	74956618	*ANKDD1B*	0.030	0.008	1.49E-04
cg16467757	8	95888145	*INTS8*	–0.022	0.006	6.92E-05
cg24590165	10	94480397	*HHEX-EXOC6*	–0.028	0.007	1.35E-04
cg05622566	11	17058241	*PLEKHA7-OR7E14P*	–0.037	0.009	6.98e-05
cg01366692	11	65291439	*SCYL1*	0.071	0.010	3.54E-12
cg25150715	12	121454808	*C12orf43*	–0.029	0.007	1.51E-05
cg25839482	15	75931953	*IMP3*	–0.020	0.004	2.54e-05

CpG, cytosine-guanine dinucleotide; SE, standard error.

Cross-trait associations of significant CpGs

Fig. 2 presents an upset plot of the MWAS results for the 111 CpGs across the 10 cardiometabolic traits, illustrating both the number of significant CpGs (with $P<4.50\times10^{-4}$ considering 111 tests) associated with each trait and the intersections among these significant CpGs across different traits. Notably, TG showed the most substantial overlap with T2DM, featuring 66 significant CpGs, followed by DBP with 52, and HDL-C with 48 significant CpGs within the set of 111. Additionally, TG, BMI, DBP, SBP, PP, and LDL-C collectively exhibited the most considerable overlap with T2DM, evidenced by 15 CpGs being significant across these traits. FG was unique in having four distinct significant CpGs that overlapped with T2DM but not with the other traits. Supplementary Table 4 contains comprehensive details of the MWAS findings.

Trans-ethnic replication in East Asians

Of the 111 significant CpGs, 56 possessed three or more instrumental meQTL, identified through the TWB. Among them, six CpGs exhibited significant MR effects in the East Asian population, achieving MR *P* values less than $\frac{0.05}{56}=5.74\times10^{-4}$ and demonstrating no evidence of pleiotropy (MR-Egger intercept $P>0.05$) or heterogeneity among instrumental variables (Cochran's Q statistic $P>0.05$). All six CpGs had F-statistics >10. Table 3 shows the six CpGs demonstrating trans-ethnic effects in East Asians based on the two-sample MR analysis. Additionally, Table 3 compares the directions of effects of these CpGs between European and East Asian populations. Notably, all six CpGs showed consistent directions of effect across both groups, underscoring their potential biological relevance to T2DM.

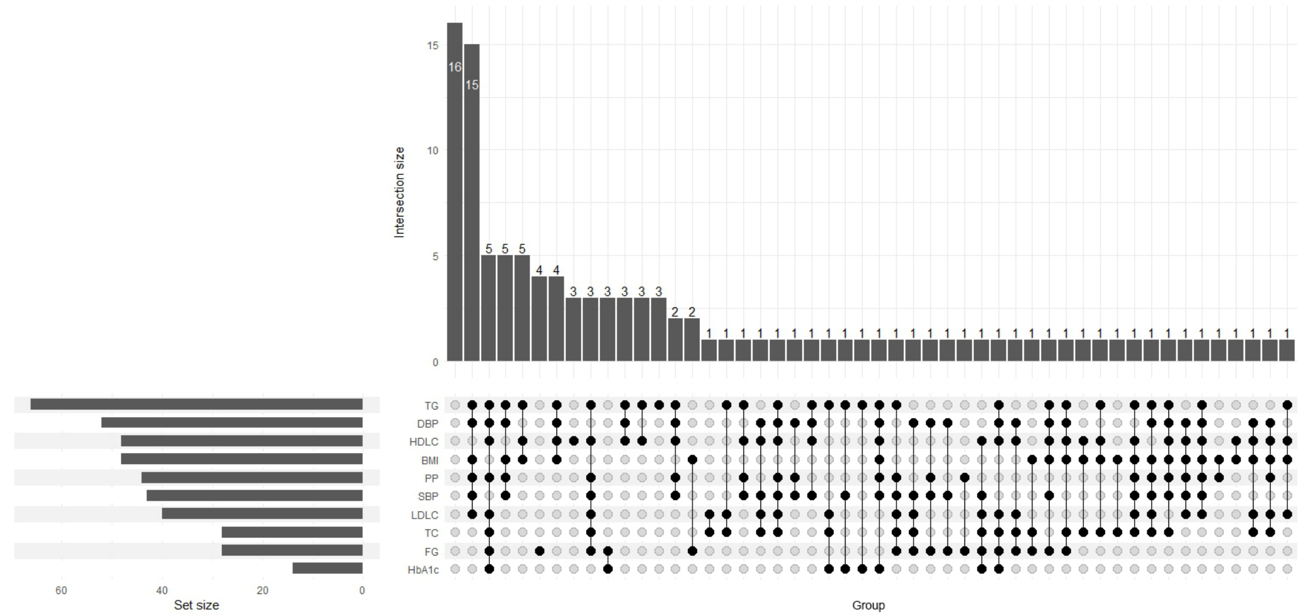


Fig. 2. Intersection analysis of methylome-wide association study (MWAS) findings for 111 cytosine-guanine dinucleotides (CpGs) across cardiometabolic traits. This figure presents an upset plot visualizing the distribution of significant CpG sites identified through MWAS among 10 cardiometabolic traits associated with type 2 diabetes mellitus (T2DM). The 'set size' section quantifies the significant CpGs associated with each individual trait. The 'intersection size' part details the count of CpGs that are significant across multiple traits. Traits associated with each significant CpG group are denoted by black dots under the 'group' category. Key traits include triglycerides (TG), body mass index (BMI), high-density lipoprotein cholesterol (HDL-C), diastolic blood pressure (DBP), systolic blood pressure (SBP), pulse pressure (PP), low-density lipoprotein cholesterol (LDL-C), fasting glucose (FG), total cholesterol (TC), and glycosylated hemoglobin (HbA1c).

Table 3. CpGs with trans-ethnic effects in East Asians

CpG	Chromosome	Position	Gene name	MR estimate (SD)	MR *P* value	Pleiotropy *P* value[a]	Heterogeneity *P* value[b]	Direction[c]	SNPs used in MR
cg00980362	1	39820405	*MACF1*	3.480 (0.366)	1.91E-21	8.21E-01	1.65E-01	+/+	3
cg24959793	5	78440068	*BHMT-JMY*	1.291 (0.265)	1.14E-06	4.36E-01	7.34E-02	+/+	18
cg16467757	8	95888145	*INTS8*	2.255 (0.600)	1.72E-04	1.38E-01	1.16E-01	+/+	4
cg13506600	9	136150361	*ABO*	1.572 (0.241)	7.21E-11	5.54E-01	9.26E-02	+/+	10
cg14271713	9	136153846	*ABO-SURF6*	−3.784 (0.577)	5.41E-11	8.63E-01	6.63E-02	-/-	7
cg06518535	17	46844976	*TTLL6*	6.880 (1.502)	4.66E-06	2.18E-01	1.86E-01	+/+	3

CpG, cytosine-guanine dinucleotide; MR, Mendelian randomization; SD, standard deviation; SNP, single-nucleotide polymorphism.
[a]The *P* value of the intercept test from MR-Egger, [b]The *P* value of the Cochran's Q statistic, [c]Directions of effects of the CpGs on type 2 diabetes mellitus in Europeans and East Asians.

eQTM and differential gene expression analyses

We assessed whether the 111 CpGs could be potential eQTM in pancreatic α- and β-cells. In α-cells, six CpGs were found to be associated with the expression of six genes ($P<0.05$). In β-cells, seven CpGs were associated with the expression of seven genes. Table 4 shows the results of differential gene expression analysis between T2DM patients and controls for the 13 genes identified by the eQTM analyses in pancreatic α- and β-cells. Among them, 2′-5′-oligoadenylate synthetase like (*OASL*) was identified as significant with a *P* value of 2.20E-03, surpassing the Bonferroni correction threshold (3.84E-03 for 13 tests) in α-cells, but no significant associations were found in β-cells. Supplementary Fig. 1 displays the residual plot of *OASL* gene expression levels in α-cells, adjusted for age and sex.

Table 4. Differential gene expression analysis results for genes regulated by eQTM

CpG				Regulated gene				eQTM statistics[a]		Differential expression *P* value[b]	Cell type
CpG	Chromosome	Position	Gene	Gene	Chromosome	Start	End	Tau	*P* value	*P* value	
cg01969012	2	27650506	*NRBP1*	*KRTCAP3*	2	27666921	27666923	0.568	2.42E-02	8.61E-01	Beta
cg22903471	2	27725779	*GCKR*	*GCKR*	2	27746304	27746306	–0.539	3.11E-02	1.44E-01	Beta
cg24072567	9	136151571	*ABO*	*SURF6*	9	136199395	136199397	–0.539	3.11E-02	9.61E-01	Beta
cg14271713	9	136153846	*ABO-SURF1*	*SURF1*	9	136218768	136218770	–0.733	2.21E-03	3.01E-01	Beta
cg23243378	11	65406311	*SIPA1*	*RELA*	11	65421849	65421851	–0.532	4.05E-02	8.34E-01	Beta
cg20447114	15	42067467	*MAPKBP1*	*PLA2G4B*	15	42140056	42140058	–0.517	4.44E-02	8.76E-01	Beta
cg17029706	17	17721645	*SREBF1*	*TOM1L2*	17	17750949	17750951	–0.822	3.57E-04	8.68E-01	Beta
cg10276098	8	8859549	*ERI1*	*ERI1*	8	8887542	8887544	0.377	4.84E-02	2.17E-01	Alpha
cg03033453	8	8902869	*ERI1*	*PPP1R3B*	8	8998304	8998306	–0.406	2.34E-02	1.21E-02	Alpha
cg02746822	8	10664645	*PINX1*	*XKR6*	8	10755462	10755464	0.367	4.25E-02	6.60E-01	Alpha
cg25150715	12	121454808	*C12orf4*	*OASL*	12	121458364	121458366	–0.368	4.22E-02	2.20E-03	Alpha
cg01268058	15	75738663	*SIN3A*	*PTPN9*	15	75761110	75761112	0.400	2.59E-02	2.92E-01	Alpha
cg22292753	17	17655165	*RAI1*	*TOM1L2*	17	17750949	17750951	–0.373	3.88E-02	5.98E-01	Alpha

eQTM, expression quantitative trait methylation; CpG, cytosine-guanine dinucleotide.
[a]The association statistics between the CpGs and the genes they regulate as eQTM, [b]The *P* values for evaluating genes differentially expressed between type 2 diabetes mellitus cases and controls using a permutation test.

Functional insights from mouse models

The findings from our database search for functional studies on the *OASL* gene in mouse models are detailed in Supplementary Table 5. We discovered that in previously published studies, mutations in the mouse Oasl2 gene were associated with reduced circulating glucose levels, indicating a potential role in glucose metabolism.

DISCUSSION

The complexity of T2DM etiology necessitates a comprehensive approach that examines the disease from multiple biological dimensions. By integrating data from different omics layers—genomics, epigenomics, and transcriptomics—our comprehensive multi-omics analysis provides significant insights into the regulatory mechanisms contributing to T2DM. This approach allows us to validate findings across multiple data sources, cross-referencing results to enhance the robustness of our conclusions.

Our results demonstrated that GWAS-identified SNPs influenced DNAm in blood samples at CpGs that could be located tens or hundreds of base pairs away. Notably, certain CpGs we identified are likely to regulate gene expression in human pancreatic α- and β-cells, and DNAm levels at the CpG cg25150715, regulated by T2DM SNPs, may lead to differential gene expression of *OASL* in α-cells, when comparing normal controls to T2DM patients. Evidence from previous studies using mouse models further corroborates the critical roles of *OASL* in glucose homeostasis, reinforcing the biological relevance of our findings.

The identification of eight novel CpGs located in or near genes not previously associated with T2DM through GWAS highlights the untapped potential of epigenetic modifications in revealing new genetic susceptibilities to metabolic diseases. For example, the presence of CpGs in genes such as BLK proto-oncogene, Src family tyrosine kinase (*BLK*) suggests that genetically regulated methylation levels may influence β-cell function and insulin secretion, as prior research has implicated *BLK* in glucose-stimulated insulin synthesis through upregulation of key pancreatic transcription factors [40]. Although Borowiec et al. [40] proposed a strong link between rare *BLK* mutations and monogenic forms of diabetes (MODY), subsequent analyses have questioned the penetrance and direct causality of these mutations [41]. Our findings indicate that methylation at CpG sites near *BLK* could serve as an additional regulatory mechanism affecting insulin secretion pathways, potentially contributing to the risk of T2DM. These epigenetic insights complement genetic studies and reinforce the com-

plexity of metabolic disease etiology, underscoring the value of integrating multi-omics approaches to better understand disease mechanisms.

Furthermore, the upset plot showing the intersections of significant CpGs among 10 cardiometabolic traits provides valuable insights into the shared epigenetic basis of these conditions, highlighting the substantial overlap of significant CpGs between T2DM and traits such as TG, BMI, blood pressure, and LDL-C. This overlap reinforces the idea that T2DM is part of a broader metabolic syndrome, with these traits being well-established risk factors. Epigenetic modifications, such as DNAm, may serve as a regulatory mechanism linking these metabolic abnormalities. For instance, alterations in DNAm may drive changes in lipid metabolism or insulin sensitivity, leading to the dysregulated glucose levels observed in T2DM. Hypomethylation of the monocyte chemoattractant protein-1 (*MCP-1*) promoter, for example, is linked to increased *MCP-1* levels, which promote inflammation, exacerbating insulin resistance and lipid metabolism dysregulation [42]. This inflammatory response may further contribute to metabolic abnormalities like hyperglycemia, dyslipidemia, and hypertension in individuals with T2DM.

The identification of six CpGs with trans-ethnic effects in East Asians suggests that these epigenetic markers may be relevant across different populations. The consistent direction of effects observed in both Europeans and East Asians enhances the generalizability of our findings between these groups and indicates that these CpGs could serve as potential biomarkers for T2DM risk in these populations. However, further studies are needed to examine these associations in additional ethnic groups to determine their broader applicability.

The unique expression of *OASL* in α-cells and its mutation-induced glucose level reduction highlight a novel link between immune mechanisms and metabolic regulation. *OASL*, through its modulation of the type I interferon (IFN) response, plays a significant role in the innate immune system's response to various pathogens, including its dual role in enhancing antiviral defenses and inhibiting autophagic mechanisms [43]. While type I IFN signaling, to which *OASL* contributes, is involved more directly in type 1 diabetes mellitus in terms of immune system regulation [44], its role in modulating the immune response suggests potential implications for T2DM as well, given the chronic inflammation and innate immune system dysfunction associated with T2DM. These discoveries underscore the potential of targeting *OASL* for diabetes management, emphasizing the need to understand gene functions across cell types for effective interventions.

While our study has uncovered significant associations and novel CpGs, it is not without limitations. The sample size for some analyses, such as the GS study and PANC-DB, was relatively small, which may affect the generalizability of those findings. Future studies with larger cohorts are necessary to validate our results and to explore the functional consequences of methylation changes in greater detail. Moreover, while the MWAS and MR associations were observed in whole blood samples, we have used human pancreatic α- and β-cells to delve into their specific functional roles in relation to the pathology of T2DM. However, investigating the effects in other tissue types, including adipose tissue, liver, and skeletal muscles, could provide further insights into the systemic nature of T2DM and its epigenetic regulation.

In conclusion, our study underscores the importance of epigenetic mechanisms in the etiology of T2DM and related cardiometabolic diseases. The identification of novel CpGs and their association with T2DM risk opens new avenues for research into the genetic and epigenetic underpinnings of metabolic diseases. This could potentially lead to the development of novel diagnostic and therapeutic strategies that target the epigenetic landscape of T2DM.

SUPPLEMENTARY MATERIALS

Supplementary materials related to this article can be found online at https://doi.org/10.4093/dmj.2025.0041.

CONFLICTS OF INTEREST

Wayne Huey-Herng Sheu has been an International editorial board members of the *Diabetes & Metabolism Journal* since 2024. He was not involved in the review process of this article. Otherwise, there was no conflict of interest.

AUTHOR CONTRIBUTIONS

Conception or design: R.H.C., C.C.W., D.D.O., W.H.H.S., H.Y.C.
Acquisition, analysis, or interpretation of data: all authors.
Drafting the work or revising: all authors.
Final approval of the manuscript: all authors.

ORCID

Ren-Hua Chung *https://orcid.org/0000-0002-9835-6333*

FUNDING

This study was supported by grants PH-112-GP-04 and PH-112-PP-10 from the National Health Research Institutes and MOST 114-2221-E-400-003-MY3 from the National Science and Technology Council in Taiwan.

ACKNOWLEDGMENTS

We thank the participants from the Taiwan Biobank. This manuscript used data acquired from the Human Pancreas Analysis Program (HPAP-RRID:SCR_016202) Database (https://hpap.pmacs.upenn.edu), a Human Islet Research Network (RRID:SCR_014393) consortium (UC4-DK-112217, U01-DK-123594, UC4-DK-112232, and U01-DK-123716).

The Taiwan Biobank data can be applied through the Taiwan Biobank at https://www.twbiobank.org.tw/. The sources for publicly available datasets used in this study are provided in Supplementary Table 1.

REFERENCES

1. DeForest N, Majithia AR. Genetics of type 2 diabetes: implications from large-scale studies. Curr Diab Rep 2022;22:227-35.
2. Tak YG, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenetics Chromatin 2015;8:57.
3. Kreitmaier P, Katsoula G, Zeggini E. Insights from multi-omics integration in complex disease primary tissues. Trends Genet 2023;39:46-58.
4. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. Philos Trans R Soc Lond B Biol Sci 2013;368:20120362.
5. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 2016;48:245-52.
6. Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating gene expression with summary association statistics to identify genes associated with 30 complex traits. Am J Hum Genet 2017;100:473-87.
7. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, et al. Opportunities and challenges for transcriptome-wide association studies. Nat Genet 2019;51:592-9.
8. Luo C, Hajkova P, Ecker JR. Dynamic DNA methylation: in the right place at the right time. Science 2018;361:1336-40.
9. Battram T, Yousefi P, Crawford G, Prince C, Sheikhali Babaei M, Sharp G, et al. The EWAS Catalog: a database of epigenome-wide association studies. Wellcome Open Res 2022;7:41.
10. Fryett JJ, Morris AP, Cordell HJ. Investigating the prediction of CpG methylation levels from SNP genotype data to help elucidate relationships between methylation, gene expression and complex traits. Genet Epidemiol 2022;46:629-43.
11. Villicana S, Castillo-Fernandez J, Hannon E, Christiansen C, Tsai PC, Maddock J, et al. Genetic impacts on DNA methylation help elucidate regulatory genomic processes. Genome Biol 2023;24:176.
12. Feng YA, Chen CY, Chen TT, Kuo PH, Hsu YH, Yang HI, et al. Taiwan Biobank: a rich biomedical research database of the Taiwanese population. Cell Genom 2022;2:100197.
13. Spracklen CN, Horikoshi M, Kim YJ, Lin K, Bragg F, Moon S, et al. Identification of type 2 diabetes loci in 433,540 East Asian individuals. Nature 2020;582:240-5.
14. Bosi E, Marchetti P, Rutter GA, Eizirik DL. Human alpha cell transcriptomic signatures of types 1 and 2 diabetes highlight disease-specific dysfunction pathways. iScience 2022;25:105056.
15. Godoy-Matos AF. The role of glucagon on type 2 diabetes at a glance. Diabetol Metab Syndr 2014;6:91.
16. Nadiger N, Veed JK, Chinya Nataraj P, Mukhopadhyay A. DNA methylation and type 2 diabetes: a systematic review. Clin Epigenetics 2024;16:67.
17. Mahajan A, Spracklen CN, Zhang W, Ng MC, Petty LE, Kitajima H, Yu et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. Nat Genet 2022;54:560-72.
18. Kurki MI, Karjalainen J, Palta P, Sipila TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature 2023;613:508-18.
19. Yu H, Du L, Yi S, Wang Q, Zhu Y, Qiu Y, et al. Epigenome-wide association study identifies Behçet's disease-associated methylation loci in Han Chinese. Rheumatology (Oxford) 2019;58:1574-84.
20. Kaestner KH, Powers AC, Naji A; HPAP Consortium; Atkinson MA. NIH initiative to improve understanding of the pancreas,

islet, and autoimmunity in type 1 diabetes: the Human Pancreas Analysis Program (HPAP). Diabetes 2019;68:1394-402.
21. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018;9:1825.
22. Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, et al. Data resource profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). Int J Epidemiol 2015;44:1181-90.
23. Porcu E, Rueger S, Lepik K; eQTLGen Consortium; BIOS Consortium; Santoni FA, et al. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. Nat Commun 2019;10:3300.
24. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658-65.
25. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-25.
26. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46:1734-9.
27. Staiger D, Stock JH. Instrumental variables regression with weak instruments. Econometrica 1997;65:557-86.
28. Hillary RF, McCartney DL, Smith HM, Bernabeu E, Gadd DA, Chybowska AD, et al. Blood-based epigenome-wide analyses of 19 common disease states: a longitudinal, population-based linked cohort study of 18,413 Scottish individuals. PLoS Med 2023;20:e1004247.
29. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res 2023;51(D1):D977-85.
30. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 2019;28:166-74.
31. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. Nat Genet 2021;53:840-60.
32. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet 2018;50:1412-25.
33. Graham SE, Clarke SL, Wu KH, Kanoni S, Zajac GJ, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature 2021;600:675-9.
34. Holter MM, Saikia M, Cummings BP. Alpha-cell paracrine signaling in the regulation of beta-cell insulin secretion. Front Endocrinol (Lausanne) 2022;13:934775.
35. Blake JA, Baldarelli R, Kadin JA, Richardson JE, Smith CL, Bult CJ, et al. Mouse Genome Database (MGD): knowledgebase for mouse-human comparative biology. Nucleic Acids Res 2021;49(D1):D981-7.
36. Baldarelli RM, Smith CM, Finger JH, Hayamizu TF, McCright IJ, Xu J, et al. The mouse Gene Expression Database (GXD): 2021 update. Nucleic Acids Res 2021;49(D1):D924-31.
37. Krupke DM, Begley DA, Sundberg JP, Richardson JE, Neuhauser SB, Bult CJ. The mouse tumor biology database: a comprehensive resource for mouse models of human cancer. Cancer Res 2017;77:e67-70.
38. Costanzo MC, von Grotthuss M, Massung J, Jang D, Caulkins L, Koesterer R, et al. The type 2 diabetes knowledge portal: an open access genetic resource dedicated to type 2 diabetes and related traits. Cell Metab 2023;35:695-710.
39. Dornbos P, Singh P, Jang DK, Mahajan A, Biddinger SB, Rotter JI, et al. Evaluating human genetic support for hypothesized metabolic disease genes. Cell Metab 2022;34:661-6.
40. Borowiec M, Liew CW, Thompson R, Boonyasrisawat W, Hu J, Mlynarski WM, et al. Mutations at the BLK locus linked to maturity onset diabetes of the young and beta-cell dysfunction. Proc Natl Acad Sci U S A 2009;106:14460-5.
41. Bonnefond A, Yengo L, Philippe J, Dechaume A, Ezzidi I, Vaillant E, et al. Reassessment of the putative role of BLK-p.A71T loss-of-function mutation in MODY and type 2 diabetes. Diabetologia 2013;56:492-6.
42. Liu ZH, Chen LL, Deng XL, Song HJ, Liao YF, Zeng TS, et al. Methylation status of CpG sites in the MCP-1 promoter is correlated to serum MCP-1 in type 2 diabetes. J Endocrinol Invest 2012;35:585-9.
43. Leisching G, Wiid I, Baker B. The association of OASL and type I interferons in the pathogenesis and survival of intracellular replicating bacterial species. Front Cell Infect Microbiol 2017;7:196.
44. Newby BN, Mathews CE. Type I interferon is a catastrophic feature of the diabetic islet microenvironment. Front Endocrinol (Lausanne) 2017;8:232.

SUPPLEMENTARY METHODS

Overview of the PrediXcan algorithm

PrediXcan is a computational method designed to understand how genetic variations influence phenotypes by exploring the underlying mechanism of gene regulation [1]. It specifically examines the impact of genetically regulated expression (GReX) on traits. GReX represents the portion of gene expression influenced by genetic factors, excluding changes driven by environmental factors or disease. Using reference transcriptome datasets from studies like Genotype-Tissue Expression (GTEx) [2], Genetic European Variation in Disease (GEUVADIS) [3], and Depression Genes and Networks (DGN) [4], PrediXcan builds predictive models to estimate GReX from genetic variants (single-nucleotide polymorphisms [SNPs]). These predictions are then tested for associations with phenotypes, reducing the number of statistical tests compared to traditional genome-wide association study (GWAS) approaches, as it evaluates gene expression rather than millions of individual SNPs. Crucially, PrediXcan doesn't require actual gene expression data from the study population, making it compatible with large genetic datasets such as the database of Genotypes and Phenotypes (dbGaP).

Overview of the S-PrediXcan algorithm

S-PrediXcan [5] is an extension of the PrediXcan method that allows for the inference of gene-trait associations using only GWAS summary statistics, eliminating the need for individual-level genotype and phenotype data. This is particularly useful when individual-level data are unavailable due to privacy concerns or study restrictions.

The core principle of S-PrediXcan is that it reconstructs the results of PrediXcan by using pre-trained weights that predict gene expression from genetic variants. These weights are derived from reference transcriptome datasets, such as GTEx and DGN, where gene expression and genetic variation have already been linked. To implement S-PrediXcan, three key inputs are required:

1. **Prediction weights:** These are coefficients from pre-trained models that predict the genetically regulated component of gene expression (GReX) based on SNP data. These models are trained using resources like GTEx, GEUVADIS, or DGN, where genetic variation is correlated with gene expression levels across individuals.
2. **SNP variance-covariance matrix:** This matrix represents the linkage disequilibrium (LD) structure between the SNPs that contribute to gene expression prediction. It ensures that the correlation between SNPs is appropriately accounted for during analysis. This information is typically derived from reference panels, such as the 1000 Genomes Project or population-specific datasets.
3. **GWAS summary statistics:** These consist of the effect sizes, standard errors, and *P* values from GWAS for each SNP being analyzed. Instead of needing individual-level data, S-PrediXcan leverages these summary statistics to infer the relationship between gene expression and the trait of interest.

Once these inputs are available, S-PrediXcan applies an analytical formula to combine the prediction weights and GWAS summary statistics, generating a predicted expression-phenotype association. The key advantage of S-PrediXcan is that it avoids the need for directly measured gene expression levels from the study population, while still capturing the genetic component of expression using the prediction weights.

Notably, the results from S-PrediXcan have been shown to closely mirror those obtained from PrediXcan, demonstrating high concordance between the two methods. This allows researchers to assess gene-level associations with traits using large-scale genome-wide association data without sacrificing power or precision.

The Taiwan Biobank cohort

Since its inception in 2008, the Taiwan Biobank (TWB) has collected samples over 150,000 individuals aged 20 and above [6]. Of these, around 129,000 individuals were genotyped using the TWB1.0 and TWB2.0 GWAS SNP arrays [7]. TWB1.0 chips, based on the customized Axiom Genome-Wide Array Plate system (Thermo Fisher Scientific, Waltham, MA, USA), comprise approximately 653,000 SNPs. TWB2.0 chips expanded upon the TWB1.0 design, encompassing about 750,000 SNPs. We adhered to the standard GWAS QC procedures, including the elimination of SNPs with Hardy-Weinberg Equilibrium $P<10^{-4}$, the removal of duplicate samples and closely related individuals, and the exclusion of participants with a mismatch between self-reported gender and genetic data-estimated biological sex. Post-QC samples were imputed via the Michigan Imputation Server, utilizing the 1000 Genomes Project phase 3 data as a reference. Subsequent analyses were based on a total of 7,369,045 SNPs, each with an imputation Rsq >0.3 and a minor allele frequency >1%.

In the Taiwan Biobank, approximately 2,300 individuals with

GWAS data underwent genome-wide methylation profiling using the Illumina Infinium MethylationEPIC BeadChip (referred to as TWB EPIC; San Diego, CA, USA). After standard QC procedures [8], 2,150 unrelated individuals and 761,617 cytosine-guanine dinucleotides (CpGs) remained. Briefly, the QC procedures was conducted using the 'bigmelon' R package (R Foundation for Statistical Computing, Vienna, Austria) [9], which involved removing outlier samples via principal component analysis, excluding samples with low bisulfite conversion rates, and normalizing the data with the 'dasen' function. Further QC steps included discarding samples with significant post-normalization changes and those with a high proportion of undetected CpG sites. After a second normalization, outlier CpG values were set to missing, and any samples or CpGs with excessive missing data were removed. Finally, probes affected by common genetic variations, cross-hybridization issues, or located on sex chromosomes were removed.

The methylation quantitative trait loci (meQTL) mapping was conducted using the software tensorQTL [10]. This tool, which utilizes graphics processing unit (GPU) acceleration, significantly enhances the speed of both cis- and trans-QTL mapping—by roughly 200 to 300 times—when compared to central processing unit (CPU)-based approaches. Our analysis focused on cis-meQTL mapping for the 111 independent and significant CpG sites identified in our European methylome-wide association study (MWAS), using samples from the Taiwan Biobank. For each CpG site, SNPs located within 1 MB were examined for their association with DNA methylation (DNAm) levels at these sites, employing a linear regression model that adjusted for age, sex, and smoking status. These SNPs were further refined using PLINK's --clump option [11], based on their association *P* values, to ensure that the LD r^2 between retained SNPs was less than 0.2. This process yielded 39,882 SNP-CpG pairs. SNPs that had association *P* values less than 1.25×10^{-6}—to adjust for the multiple tests performed across the 39,882 pairs—were identified as meQTLs for the 111 CpG sites. In total, 685 meQTL were discovered, and their association statistics were subsequently utilized in Mendelian randomization analyses focused on the East Asian population.

The PANC-DB dataset

PANC-DB serves as an open-source repository designed for the storage, dissemination, and visualization of cellular and molecular data produced by the Human Pancreas Analysis Program (HPAP), thereby enhancing the accessibility of genomic and islet function data for the diabetes research community [12]. It integrates three main facets: data management and access, computational biology and data science, and communication and outreach efforts. HPAP, a key component of the Human Islet Research Network (HIRN), focuses on the detailed phenotyping of the human endocrine pancreas to uncover the processes leading to beta-cell loss or dysfunction in both type 1 diabetes mellitus and T2DM. We acquired DNAm data, derived from whole-genome bisulfite sequencing, and gene expression data, obtained through RNA sequencing (RNA-Seq) of pancreatic α- and β-cells, from PANC-DB. Analysis of the DNAm data revealed DNAm levels at 82 of the 87 CpGs identified in our MWAS, which were subsequently analyzed.

For the RNA-Seq data, we adhered to the mRNA Analysis Pipeline (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline/) from the National Cancer Institute Genomic Data Commons (GDC). This pipeline begins with an Alignment Workflow that uses the Spliced Transcripts Alignment to a Reference (STAR) tool to implement a two-pass alignment method. Each read group is aligned separately by STAR, then merged to produce a comprehensive alignment, incorporating a splice junction detection step as recommended by the International Cancer Genome Consortium (ICGC). The result is a genomic binary alignment/map (BAM) file containing both aligned and unaligned reads. QC is rigorously applied before alignment with FastQC and after alignment with Picard Tools. Finally, featureCounts [13] is utilized to generate raw counts, from which transcript per million (TPM) values are calculated for gene expression analysis.

In our study, we conducted a differential gene expression analysis by comparing the mean TPM values between T2DM patients and control individuals without a history of diabetes. We first adjusted the TPM for each gene by age and sex using a linear regression model, then derived residuals for further analysis. Given the limited size of the sample, we employed a permutation test to assess the association between these residuals and T2DM status. Specifically, we calculated the difference in mean TPM values between the T2DM and control groups. In the permutation test, we randomly shuffled the diabetes status among the samples and recalculated the mean TPM difference for each permutation. This process was repeated 10,000 times. The permutation *P* value was determined by the fraction of permutations where the absolute mean difference was greater than the absolute observed mean difference.

SUPPLEMENTARY REFERENCES

1. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015;47:1091-8.
2. Carithers LJ, Moore HM. The genotype-tissue expression (GTEx) project. Biopreserv Biobank 2015;13:307-8.
3. Lappalainen T, Sammeth M, Friedlander MR, 't Hoen PA, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013;501:506-11.
4. Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res 2014;24:14-24.
5. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018;9:1825.
6. Fan CT, Lin JC, Lee CH. Taiwan Biobank: a project aiming to aid Taiwan's transition into a biomedical island. Pharmacogenomics 2008;9:235-46.
7. Feng YA, Chen CY, Chen TT, Kuo PH, Hsu YH, Yang HI, et al. Taiwan Biobank: a rich biomedical research database of the Taiwanese population. Cell Genom 2022;2:100197.
8. Chung RH, Chiu YF, Wang WC, Hwu CM, Hung YJ, Lee IT, et al. Multi-omics analysis identifies CpGs near G6PC2 mediating the effects of genetic variants on fasting glucose. Diabetologia 2021;64:1613-25.
9. Gorrie-Stone TJ, Smart MC, Saffari A, Malki K, Hannon E, Burrage J, et al. Bigmelon: tools for analysing large DNA methylation datasets. Bioinformatics 2019;35:981-6.
10. Taylor-Weiner A, Aguet F, Haradhvala NJ, Gosai S, Anand S, Kim J, et al. Scaling computational genomics to millions of individuals with GPUs. Genome Biol 2019;20:228.
11. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
12. Kaestner KH, Powers AC, Naji A, Atkinson MA. NIH initiative to improve understanding of the pancreas, islet, and autoimmunity in type 1 diabetes: the Human Pancreas Analysis Program (HPAP). Diabetes 2019;68:1394-402.
13. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923-30.

Supplementary Table 1. Datasets/models used in our analysis

Dataset/Model	Type	Sample size	Ethnicity	Purpose	Reference
Understanding Society: UK Household Longitudinal Study	DNAm prediction model trained on GWAS and DNAm data	1,120	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/35930604/
DIAMANTE	GWAS summary statistics	80,154 T2DM cases, 853,816 controls	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/30297969/
FinnGEN	GWAS summary statistics	38,657 T2DM cases, 310,131 controls	European	MR	https://r9.finngen.fi/
BMI GWAS	GWAS summary statistics	806,834 with BMI measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/30239722/
Fasting glucose GWAS	GWAS summary statistics	200,622 with fasting glucose measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34059833/
HbA1c GWAS	GWAS summary statistics	146,806 with HbA1c measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34059833/
Systolic blood pressure (SBP) GWAS	GWAS summary statistics	757,601 with SBP measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/30224653/
Diastolic blood pressure (DBP) GWAS	GWAS summary statistics	757,601 with DBP measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/30224653/
Pulse pressure (PP) GWAS	GWAS summary statistics	757,601 with PP measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/30224653/
Total cholesterol GWAS	GWAS summary statistics	1,320,016 with total cholesterol measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34887591/
Triglycerides (TG) GWAS	GWAS summary statistics	1,320,016 with TG measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34887591/
LDL-C GWAS	GWAS summary statistics	1,320,016 with LDL-C measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34887591/
HDL-C GWAS	GWAS summary statistics	1,320,016 with HDL-C measurements	European	MWAS	https://pubmed.ncbi.nlm.nih.gov/34887591/
AGEN GWAS	GWAS summary statistics	77,418 T2DM cases, 356,122 controls	East Asian	MR	https://pubmed.ncbi.nlm.nih.gov/32499647/
Taiwan Biobank	Methylation and GWAS chip	2,150 samples	East Asian	MR	https://www.twbiobank.org.tw/
EWAS Catalog	EWAS summary statistics	757 incident cases, 16,992 controls	European	Validation	https://pubmed.ncbi.nlm.nih.gov/35592546/
MeQTL EPIC	meQTL database	2,358 samples	European	MR	https://pubmed.ncbi.nlm.nih.gov/37525248/
PANC-DB	WGBS and RNA-Seq data in human pancreatic α- and β-cells	eQTM analysis: 17 and 10 α- and β-cell samples, respectively. Differential gene expression analysis: 6 T2DM cases and 26 controls for α-cells, 6 T2DM cases and 22 controls for β-cells	European	eQTM and differential gene expression analyses	https://pubmed.ncbi.nlm.nih.gov/31127054/

DNAm, DNA methylation; GWAS, genome-wide association study; MWAS, methylome-wide association study; DIAMANTE, Diabetes Meta-analysis of Trans-ethnic Association Studies; T2DM, type 2 diabetes mellitus; MR, Mendelian randomization; BMI, body mass index; HbA1c, glycosylated hemoglobin; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; AGEN, Asian Genetic Epidemiology Network; EWAS, epigenome-wide association study; MeQTL, methylation quantitative trait loci; PANC-DB, Data portal of The Human Pancreas Analysis Program; WGBS, whole-genome bisulfite sequencing; eQTM, expression quantitative trait methylation.

Supplementary Table 2. Summary statistics for significant and independent CpGs identified by MWAS and MR for type 2 diabetes mellitus
https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fe-dmj.org%2Fupload%2Fmedia%2Fdmj-2025-0041-Supplementary-Table-2.xlsx&wdOrigin=BROWSELINK

Supplementary Table 3. Lead SNP information from the MWAS and MR analyses

CpG	Chromosome	Position	MWAS				MR			
			Lead SNP	DIAMANTE *P* value	Weight[a]	Distance[b]	Lead SNP	FinnGen *P* value	MeQTL *P* value	Distance[b]
cg03179008	1	39620300	rs3768301	2.81E-25	−4.39E-05	−250,493	rs16837533	1.21E-08	9.23E-117	83,547
cg06697744	1	39680237	rs16837533	1.03E-19	2.31E-03	−23,610	rs1775654	9.57E-08	6.44E-28	269,843
cg08719237	1	39728532	rs16826069	2.81E-25	−1.10E-03	−68,523	rs56809193	1.51E-08	3.01E-62	−119,234
cg00980362	1	39820405	rs61779275	1.07E-25	−3.77E-04	95	rs16837533	1.21E-08	1.76E-40	−116,558
cg21619813	1	219679937	rs2820441	8.56E-16	−2.53E-04	−55,023	rs59584436	6.00E-09	4.51E-34	−15,309
cg06162668	2	638091	rs1320330	5.72E-09	4.54E-04	15,866	rs4854331	1.20E-04	1.06E-43	−20,470
cg01969012	2	27650506	rs780094	9.59E-22	2.90E-04	−90,731	rs7564625	2.42E-03	3.06E-39	211,891
cg05102552	2	27650867	rs780110	1.67E-11	2.91E-03	−34,521	rs4665986	2.66E-04	5.39E-73	104,301
cg22903471	2	27725779	rs3845687	1.68E-08	8.56E-04	35,880	rs3749147	5.27E-04	6.67E-25	126,139
cg10337841	2	27747795	rs780110	1.67E-11	6.31E-04	62,407	rs780105	8.27E-06	8.93E-46	−69,324
cg16965262	2	43633292	rs10186921	1.70E-07	1.58E-03	−11,264	rs6544654	3.65E-05	5.89E-18	276
cg05010058	2	65284262	rs2540951	1.73E-15	−3.77E-04	7,526	rs12468907	7.85E-03	7.94E-11	110,230
cg05293897	2	226913713	rs2713538	2.51E-24	1.17E-05	−220,803	rs60271363	1.75E-09	3.96E-40	−13,391
cg04891645	2	227050097	rs10439362	3.28E-34	−5.02E-03	23,361	rs952227	1.76E-19	1.12E-60	11,983
cg10397322	3	23244051	rs1496653	2.49E-17	−9.56E-04	−210,739	rs1496653	1.81E-10	2.37E-65	210,739
cg18257541	3	23244062	rs1496653	2.49E-17	−9.24E-04	−210,728	rs11926494	1.10E-06	5.83E-22	14,552
cg10632094	3	23244068	rs1496653	2.49E-17	−7.74E-04	−210,722	rs1908475	1.10E-06	5.00E-18	24,088
cg24311743	3	23244116	rs9863600	1.51E-10	−5.61E-07	53,284	rs4422335	1.40E-05	1.45E-09	−56,709
cg16258503	3	63850278	rs704362	2.13E-11	−2.77E-04	−28,062	rs11922435	1.33E-04	1.18E-04	140,529
cg16021135	3	123130669	rs6798189	5.28E-26	2.17E-03	35,357	rs7612096	1.48E-05	1.54E-84	5,579
cg12781915	3	185544099	rs6444082	3.56E-35	2.25E-03	7,876	rs28485641	1.47E-08	2.89E-15	−132,483
cg01899937	4	185723505	rs12504993	9.67E-14	1.68E-05	5,019	rs6552828	1.03E-07	8.36E-12	1,911
cg20885688	5	74956618	rs6864091	7.46E-13	−1.05E-03	−53,384	rs6881648	9.75E-07	3.06E-100	35,231
cg27359689	5	75034394	rs42302	6.66E-12	−1.05E-03	68,840	rs7715806	1.66E-04	8.58E-171	−7
cg17746515	5	78439992	rs2364594	1.69E-10	7.73E-04	−13,907	rs35893235	3.49E-04	8.34E-62	154,825
cg24959793	5	78440068	rs2121107	2.05E-10	1.00E-03	−27,936	rs1717565	4.67E-04	4.11E-108	301
cg19277069	5	102089501	rs258244	2.42E-08	−1.12E-04	−72,250	rs58786975	9.67E-05	2.17E-04	−511,391
cg26671988	5	102089760	rs17296280	8.02E-11	4.22E-03	−270,987	rs58786975	9.67E-05	3.25E-11	−511,650
cg12989544	6	153429875	rs9371676	2.93E-07	−3.14E-03	−16,090	rs9397581	7.32E-04	8.82E-64	−46,285
cg16837680	8	8084974	rs2921077	9.14E-09	−3.69E-04	−219,528	rs2976940	7.88E-03	2.60E-31	194,488
cg14469574	8	8365789	rs2288673	2.07E-09	2.78E-04	−494,487	rs2979208	3.03E-04	2.08E-23	−179,622
cg20987067	8	8747297	rs2956244	5.36E-10	−7.63E-04	−137,869	rs399123	8.27E-04	3.73E-52	240
cg10276098	8	8859549	rs2921383	4.41E-10	4.64E-03	−32,672	rs587328	1.46E-05	1.25E-05	950,173
cg03033453	8	8902869	rs7005133	3.02E-09	−3.41E-04	1,647	rs6601299	3.23E-05	2.06E-12	281,822
cg13288100	8	8915148	rs7853	9.62E-10	−2.03E-04	24,334	rs2979256	2.38E-04	4.41E-58	−43,438
cg00184457	8	8946301	rs2921383	4.41E-10	−3.59E-04	54,080	rs2979241	5.54E-04	5.48E-14	−642,948
cg05566310	8	9807662	rs60384372	7.88E-13	−3.12E-04	−166,922	rs3750310	1.39E-05	6.42E-30	475,764
cg03502219	8	9978949	rs17689007	5.62E-13	−1.02E-03	4,125	rs28712068	2.77E-05	3.95E-34	272,196
cg01048752	8	10001523	rs35013996	3.73E-08	4.18E-03	54,378	rs35437983	4.52E-03	2.45E-22	3,530
cg07426455	8	10199561	rs17151859	7.07E-08	1.47E-07	−38,419	rs28406568	1.03E-04	1.05E-33	−6,337

(*Continued to the next page*)

Supplementary Table 3. Continued

CpG	Chromosome	Position	MWAS				MR			
			Lead SNP	DIAMANTE *P* value	Weight[a]	Distance[b]	Lead SNP	FinnGen *P* value	MeQTL *P* value	Distance[b]
cg01236304	8	10614369	rs28570522	1.26E-10	−1.95E-03	−16,199	rs73203359	1.16E-04	5.26E-06	−312,386
cg08673857	8	10634343	rs11250078	1.86E-08	9.45E-05	−32,238	rs1876511	5.68E-03	1.02E-04	−755,060
cg02746822	8	10664645	rs7350075	2.75E-08	2.59E-03	39,646	rs615632	2.72E-04	4.92E-30	−868,324
cg06325495	8	10767745	rs7016385	4.31E-11	1.47E-03	−11,727	rs4841282	2.61E-04	1.30E-48	−792,141
cg18477422	8	10773493	rs4841457	2.84E-11	−1.12E-02	−18	rs34741518	4.18E-05	3.44E-36	3,367
cg14919250	8	10917293	rs4841486	8.03E-07	5.77E-04	5,804	rs11986122	1.24E-03	1.02E-26	−907,344
cg21725652	8	10918152	rs2409685	4.83E-07	3.59E-04	8,959	rs56159799	2.12E-03	2.74E-06	−596,398
cg18317842	8	10921502	rs4642600	4.43E-08	−9.99E-05	−91,523	rs11986122	1.24E-03	3.79E-18	−911,553
cg09697033	8	10929599	rs11783749	5.87E-11	5.20E-04	127,544	rs11986122	1.24E-03	1.80E-14	−919,650
cg25014118	8	10948462	rs4841495	2.20E-10	4.77E-04	−10,362	rs4841660	1.59E-03	4.07E-25	879,841
cg17016032	8	11020264	rs17782536	5.43E-09	9.73E-04	−1,418	rs17725534	8.17E-04	2.28E-48	−10,595
cg10813958	8	11247920	rs6980481	1.00E-08	−8.38E-03	3,079	rs2409716	3.50E-04	8.02E-49	−237,172
cg02034205	8	11382291	rs2618443	9.82E-08	−9.92E-04	−2,265	rs28380506	6.61E-03	9.46E-12	−926,523
cg20677018	8	41508391	rs13262861	3.04E-26	8.71E-03	−186	rs2977864	6.46E-07	9.77E-14	−12,077
cg09232650	8	41523757	rs508419	5.43E-27	−1.08E-02	766	rs13254454	8.49E-03	1.34E-46	−22,195
cg16467757	8	95888145	rs12548874	3.17E-11	2.19E-03	−47,019	rs6991067	5.36E-06	1.68E-55	−12,469
cg22283921	8	95961819	rs10097617	2.44E-14	−3.04E-03	193	rs2515140	9.31E-05	2.44E-15	−649,477
cg21160290	9	136149941	rs505922	3.65E-12	−6.25E-03	712	rs9411475	1.43E-04	1.29E-17	−22,673
cg13506600	9	136150361	rs505922	3.65E-12	−2.60E-03	1,132	rs633862	1.92E-09	1.84E-62	5,083
cg12855781	9	136150585	rs657152	2.75E-10	−4.98E-04	11,320	rs630014	9.26E-05	1.34E-13	−863
cg11698867	9	136150802	rs582118	6.00E-12	−2.25E-04	5,331	rs630014	9.26E-05	1.12E-24	−1,080
cg07241568	9	136150823	rs582118	6.00E-12	−2.25E-04	5,352	rs635634	1.09E-06	1.14E-56	4,177
cg24072567	9	136151571	rs582118	6.00E-12	−2.36E-03	6,100	rs1179011	9.12E-06	9.36E-24	34,812
cg14271713	9	136153846	rs505922	3.65E-12	4.90E-04	4,617	rs630014	9.26E-05	6.98E-23	−4,124
cg05289649	10	80925798	rs1613299	2.32E-26	1.00E-02	−28,987	rs35339173	1.58E-05	6.34E-238	−20,879
cg21213154	10	80927984	rs10824724	2.22E-10	−1.48E-03	4,102	rs703994	3.34E-08	5.70E-26	−1,948
cg24340021	10	94304119	rs3118967	1.00E-12	−3.88E-05	95,372	rs201503938	1.61E-04	1.43E-33	−89,105
cg24590165	10	94480397	rs1832886	6.67E-61	1.72E-03	2,858	rs7903302	5.22E-18	2.09E-49	−50,886
cg00782718	11	2197180	rs4929965	9.67E-23	1.02E-03	−106	rs11042978	2.22E-04	1.25E-34	1,238
cg05622566	11	17058241	rs10832731	7.71E-09	−1.94E-03	−39,454	rs739689	3.33E-03	6.66E-06	359,263
cg01366692	11	65291439	rs11601767	2.49E-12	−6.53E-04	−24,943	rs1783541	5.49E-03	9.91E-08	3,360
cg23243378	11	65406311	rs1075692	5.42E-09	3.94E-04	193,864	rs78712280	1.78E-03	2.15E-35	157,192
cg14286243	11	68867220	rs10896415	1.69E-07	8.52E-03	27,886	rs3750971	2.19E-05	8.34E-189	−36,513
cg25903143	11	68920466	rs1060435	2.35E-07	6.11E-03	64,871	rs10896402	2.49E-04	3.93E-61	−94,149
cg14036981	11	68920648	rs1060435	2.35E-07	8.86E-03	65,053	rs10896402	2.49E-04	2.78E-115	−94,331
cg25179853	11	68924577	rs1060435	2.35E-07	1.08E-04	68,982	rs72932523	1.31E-05	1.23E-48	−18,457
cg23740940	11	68924746	rs1060435	2.35E-07	6.51E-03	69,151	rs61881115	3.38E-03	2.22E-16	72,479
cg00048149	12	26472528	rs11048458	3.01E-10	1.49E-03	6,943	rs10842716	9.17E-05	9.24E-35	26,959
cg07742264	12	121419635	rs2708081	1.24E-09	−2.39E-03	−43,653	rs1169309	8.72E-12	3.41E-09	19,557
cg25150715	12	121454808	rs1169302	5.37E-16	−7.19E-04	22,506	rs1169302	2.96E-13	1.00E-12	−22,506

(*Continued to the next page*)

Supplementary Table 3. Continued

CpG	Chromosome	Position	MWAS				MR			
			Lead SNP	DIAMANTE *P* value	Weight[a]	Distance[b]	Lead SNP	FinnGen *P* value	MeQTL *P* value	Distance[b]
cg00090265	12	133140305	rs11147192	2.39E-07	−9.04E-05	79,756	rs11147192	2.22E-03	1.61E-63	−79,756
cg27454283	13	51095835	rs9316500	3.45E-08	−4.78E-03	1,721	rs2182784	1.34E-04	2.42E-19	37,586
cg21805149	14	91871665	rs10134560	1.32E-08	1.44E-04	−58,487	rs7146002	6.53E-04	9.19E-39	7
cg16461996	15	41822953	rs2412619	4.75E-12	9.88E-04	−3,044	rs1200352	2.27E-03	4.69E-36	8,423
cg08686008	15	41876221	rs1023193	2.30E-08	−7.01E-03	20,485	rs28506183	9.32E-03	1.27E-75	310,724
cg20447114	15	42067467	rs4924576	3.13E-09	1.16E-04	−52,126	rs2616	9.55E-04	2.11E-18	−21,807
cg14158153	15	42141147	rs2297378	1.85E-07	9.79E-05	324,957	rs13380315	1.09E-03	4.41E-20	−20,710
cg01268058	15	75738663	rs35134156	3.88E-11	1.05E-03	−1,576,769	rs11638974	1.53E-04	1.02E-39	60,643
cg25839482	15	75931953	rs13737	1.77E-10	−4.51E-04	−176	rs6495182	7.14E-05	5.12E-20	−117,565
cg12679965	15	75932781	rs13737	1.77E-10	−2.60E-03	652	rs11638974	1.53E-04	1.13E-18	−133,475
cg23627990	15	77364890	rs35134156	3.88E-11	6.86E-04	49,458	rs62008430	9.65E-09	4.85E-32	202,936
cg08075384	15	77779509	rs11629835	2.03E-13	−1.14E-03	−7,960	rs7119	1.22E-09	5.31E-12	−1,877
cg02931208	15	77866341	rs12593111	1.80E-17	4.55E-03	12,667	rs2682922	6.09E-08	1.15E-25	−1,335
cg10846857	15	77907603	rs4886889	2.83E-11	1.14E-03	5,256	rs11632729	2.16E-08	4.06E-39	−185,597
cg01959774	16	290115	rs6600204	4.95E-12	−2.65E-03	−10,062	rs7189326	2.50E-03	2.28E-28	23,291
cg01474600	16	315564	rs6600204	4.95E-12	−6.69E-04	15,387	rs7189326	2.50E-03	1.08E-150	−2,158
cg03598338	16	75300041	rs8054218	7.22E-20	6.32E-04	47,880	rs55993634	8.42E-14	1.60E-73	−63,278
cg04987608	16	75300043	rs7202877	9.75E-20	1.50E-04	52,798	rs8046145	2.38E-07	2.00E-76	−5,022
cg18059928	16	75300238	rs8056814	7.60E-25	6.97E-04	47,911	rs3743614	2.60E-07	1.45E-34	−19,280
cg15598217	17	4001879	rs8068804	1.31E-13	−1.88E-03	16,015	rs2074991	2.15E-04	6.09E-10	−163,253
cg05511978	17	4092306	rs17763551	2.51E-13	−1.12E-03	209,997	rs35050090	8.96E-09	2.95E-43	−44,088
cg22292753	17	17655165	rs11655029	2.08E-11	−1.55E-03	5,993	rs117908897	7.86E-04	9.14E-48	−70,019
cg17029706	17	17721645	rs12946913	1.14E-10	3.01E-04	15,313	rs8076747	1.79E-05	5.46E-73	131,950
cg22116239	17	17741676	rs1108646	6.85E-11	−2.41E-03	−9,802	rs75035298	2.28E-04	3.97E-59	−268,662
cg06518535	17	46844976	rs35895680	1.08E-15	3.35E-04	−215,346	rs7217007	1.34E-07	1.06E-35	−279
cg08228953	19	7979674	rs3679	4.61E-12	−4.40E-03	1,244	rs2303700	2.29E-08	2.55E-44	−3,145
cg02942825	19	46185348	rs112128231	4.49E-09	−1.19E-03	−3,009	rs34642101	3.41E-07	1.13E-37	−21,358
cg09118261	20	32568722	rs4911414	3.33E-10	2.28E-04	−160,722	rs6141433	3.17E-04	7.58E-47	−19,860
cg17113955	20	32610490	rs1007090	1.01E-10	6.76E-04	27,619	rs6059685	2.51E-05	9.23E-41	121,167
cg03567001	20	32668840	rs1007090	1.01E-10	−2.00E-04	85,969	rs2747568	1.44E-03	3.22E-14	−275,840
cg17746527	20	48834262	rs11699802	8.84E-12	4.74E-04	2,127	rs4253439	2.15E-07	2.88E-46	−26,251

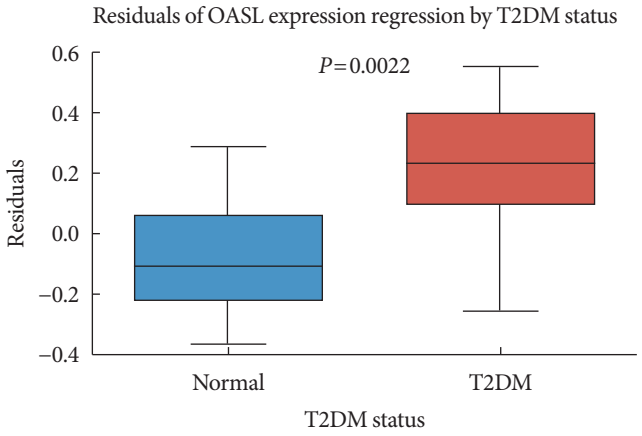
SNP, single-nucleotide polymorphism; MWAS, methylome-wide association study; MR, Mendelian randomization; CpG, cytosine-guanine dinucleotide; DIAMANTE, Diabetes Meta-analysis of Trans-ethnic Association Studies; MeQTL, methylation quantitative trait loci.
[a]Weight of the SNP in the DNA methylation prediction model, [b]Distance of the SNP to the CpG: A negative value indicates that the SNP is upstream of the CpG, while a positive value indicates that the SNP is downstream of the CpG.

Supplementary Table 4. MWAS results for the 10 cardiometabolic traits
https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fe-dmj.org%2Fupload%2Fmedia%2Fdmj-2025-0041-Supplementary-Table-4.xlsx&wdOrigin=BROWSELINK

Supplementary Table 5. Functional analysis of the differentially expressed genes in mouse models

Gene	Allelic composition	Consequence of mutant allele	Genetic background	Genotype ID	Annotated term	Phenotype summary category	Reference
Oasl1							
	Oasl1<tm1b(EUCOMM)Hmgu>/Oasl1<tm1b(EUCOMM)Hmgu>	Oasl1<tm1b(EUCOMM)Hmgu>: Null/knockout	C57BL/6N-Oasl1<tm1b(EUCOMM)Hmgu>/BayMmucd	MGI:6461494	Preweaning lethality, incomplete penetrance[a]	Mortality/aging	J:211773
	Oasl1<tm1Yjk>/Oasl1<tm1Yjk>	Oasl1<tm1Yjk>: Null/knockout	B6.129P2-Oasl1<tm1Yjk>	MGI:5498951	Abnormal circulating interferon level	Homeostasis/metabolism	J:194467
	Oasl1<tm1Yjk>/Oasl1<tm1Yjk>	Oasl1<tm1Yjk>: Null/knockout	B6.129P2-Oasl1<tm1Yjk>	MGI:5498951	Abnormal circulating interferon level	Immune system	J:194467
	Oasl1<tm1Yjk>/Oasl1<tm1Yjk>	Oasl1<tm1Yjk>: Null/knockout	B6.129P2-Oasl1<tm1Yjk>	MGI:5498951	Decreased susceptibility to Picornaviridae infection	Immune system	J:194467
	Oasl1<tm1Yjk>/Oasl1<tm1Yjk>	Oasl1<tm1Yjk>: Null/knockout	B6.129P2-Oasl1<tm1Yjk>	MGI:5498951	Decreased susceptibility to Picornaviridae infection induced morbidity/mortality	Immune system	J:194467
	Oasl1<tm1Yjk>/Oasl1<tm1Yjk>	Oasl1<tm1Yjk>: Null/knockout	B6.129P2-Oasl1<tm1Yjk>	MGI:5498951	Decreased susceptibility to Picornaviridae infection induced morbidity/mortality	Mortality/aging	J:194467
Oasl2							
	Oasl2<tm1a(EUCOMM)Wtsi>/Oasl2<tm1a(EUCOMM)Wtsi>	Oasl2<tm1a(EUCOMM)Wtsi>: Null/knockout	C57BL/6N-Oasl2<tm1a(EUCOMM)Wtsi>/Wtsi	MGI:5781689	Decreased circulating glucose level[a]	Homeostasis/metabolism	J:175295

[a]Phenotypes found in the International Mouse Phenotyping Consortium (IMPC).



Supplementary Fig. 1. Box plot illustrating the residuals of 2′-5′-oligoadenylate synthetase like (*OASL*) gene expression, adjusted for age and sex, in pancreatic alpha cells across the normal and type 2 diabetes mellitus (T2DM) groups. The displayed *P* value signifies the difference in means between the two groups, determined through a permutation test.