



Brief Report

Shared Genetics between Age at Menarche and Type 2 Diabetes Mellitus: Genome-Wide Genetic Correlation Study

Yuan-Fang Cheng ¹, Cheng-Yi Yang ² and Meng-Che Tsai ^{3,4,5,*} ¹ School of Medicine, College of Medicine, National Cheng Kung University, Tainan 70101, Taiwan² Department of Statistics, College of Management, National Cheng Kung University, Tainan 70101, Taiwan³ Department of Pediatrics, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, 138 Shengli Road, Tainan 70403, Taiwan⁴ Department of Genomic Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan 70101, Taiwan⁵ Department of Medical Humanities and Social Medicine, College of Medicine, National Cheng Kung University, Tainan 70101, Taiwan

* Correspondence: ache93@yahoo.com.tw; Tel.: +886-6-2353535-4190

Abstract: Background: Age at menarche (AAM) has been associated with type 2 diabetes mellitus (T2DM). However, little is known about their shared heritability. **Methods:** Our data comes from the Taiwan Biobank. Genome-wide association studies (GWASs) were conducted to identify single-nucleotide polymorphisms (SNPs) related to AAM-, T2DM-, and T2DM-related phenotypes, such as body fat percentage (BFP), fasting blood glucose (FBG), and hemoglobin A1C (HbA_{1C}). Further, the conditional false discovery rate (cFDR) method was applied to examine the shared genetic signals. **Results:** Conditioning on AAM, Quantile-quantile plots showed an earlier departure from the diagonal line among SNPs associated with BFP and FBG, indicating pleiotropic enrichments among AAM and these traits. Further, the cFDR analysis found 39 independent pleiotropic loci that may underlie the AAM-T2DM association. Among them, *FN3KRP* rs1046896 (cFDR = 6.84×10^{-49}), *CDKAL1* rs2206734 (cFDR = 6.48×10^{-10}), *B3GNTL1* rs58431774 (cFDR = 2.95×10^{-10}), *G6PC2* rs1402837 (cFDR = 1.82×10^{-8}), and *KCNQ1* rs60808706 (cFDR = 9.49×10^{-8}) were highlighted for their significant genetic enrichment. The protein–protein interaction analysis revealed a significantly enriched network among novel discovered genes that were mostly found to be involved in the insulin and glucagon signaling pathways. **Conclusions:** Our study highlights potential pleiotropic effects across AAM and T2DM. This may shed light on identifying the genetic causes of T2DM.

Keywords: age at menarche; type 2 diabetes mellitus; genome-wide association study; conditional false discovery rate; Taiwan Biobank



Citation: Cheng, Y.-F.; Yang, C.-Y.; Tsai, M.-C. Shared Genetics between Age at Menarche and Type 2 Diabetes Mellitus: Genome-Wide Genetic Correlation Study. *Biomedicines* **2024**, *12*, 157. <https://doi.org/10.3390/biomedicines12010157>

Academic Editor: Guanglin Zhang

Received: 19 December 2023

Revised: 7 January 2024

Accepted: 9 January 2024

Published: 11 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the International Diabetes Federation, the global number of adults living with type 2 diabetes mellitus (T2DM) is predicted to rise to 643 million by 2030 and 783 million by 2045 [1]. Moreover, Taiwan has the highest prevalence of diabetes among Asian countries, and the prevalence is still rapidly rising by 25,000 people per year [2]. T2DM, with its occurrence mainly due to excessive body adiposity and insulin resistance, can lead to severe chronic complications, including cardiovascular, central and peripheral nervous, and immune disorders. Studies of twins have shown that genetics plays an essential role in the development of T2DM [3]. According to prior studies, several important genes have been indicated, including *CAPN10* (calpain 10), *PPARG* (peroxisome proliferator-activated receptor gamma), *KCNJ11* (potassium inwardly rectifying channel subfamily J member 11), *TCF7L2* (transcription factor 7-like 2), *KCNQ1* (potassium voltage-gated channel subfamily Q member 1), and *MTNRB1* (melatonin receptor B1) [4].

Early age at menarche (AAM) is usually defined as menarche before the age of 12 years. Over the past 30 years, AAM has decreased in Taiwanese children, with a rate of decline of approximately 0.43 years per decade [5]. Early menarche has been associated with physical and psychosocial problems, including anxiety/depression, substance use, suicidal behaviors, early fusion of epiphyseal growth plates, cardiovascular diseases, and metabolic syndromes [6]. AAM is a hereditary trait, but its known genetic determinants only explain part of the variance [7].

Early AAM and T2DM are highly heritable traits with rising prevalence, and both may adversely affect healthy life outcomes [8,9]. Studies have revealed an association between AAM and the risk of T2DM [10], as women with early menarche have a higher risk of T2DM in adulthood. Although excessive adiposity is assumed to be a common driver, the two traits may share genetic loci, but their shared heritability remains unclear. Therefore, a genome-wide association study (GWAS) is required to investigate whether AAM and T2DM have common heritable origins and further determine the underlying molecular pathways.

In this study, we aimed to identify shared genetic variants of AAM and T2DM and investigate the connection between the two traits using a nationally representative database. However, GWAS has underlying shortcomings, such as low genetic coverage, missing phenotypes, and low allele frequency. Quality control (QC) of genotypic and phenotypic data was conducted to reduce biases in GWAS. We hypothesize that genetic variants may show differential associations with T2DM on the condition of their associations with AAM.

2. Methods

2.1. Participants

Genotypic and phenotypic data were available for 88,347 participants from the Taiwan Biobank (TWB, <https://www.twbiobank.org.tw>, accessed on 27 April 2023). The TWB is a government-supported, prospective cohort study with a wide range of phenotype and genome data for the Taiwanese population [11]. TWB 2.0 axiom genome-wide array plate was designed for whole-genome genotyping. There are approximately 750,000 single-nucleotide polymorphisms (SNPs) in TWB 2.0 covering specific SNPs associated with diseases, drug metabolism, and drug response in the Taiwanese population [12]. Phenotypic data were obtained from questionnaires and measurements [13]. Participants were asked to report their age, education level, age at enrollment, AAM, diagnosis of diabetes, types of diabetes, and family history of diabetes in the questionnaire. Other T2DM-related phenotypes of interest, including body fat percentage (BFP), fasting blood glucose (FBG), and hemoglobin A1C (HbA_{1C}), were obtained from measurements [14,15]. The present study was approved by the Institutional Review Board of National Cheng Kung University.

2.2. Quality Control Procedures

QC procedures were implemented using PLINK (<https://www.cog-genomics.org/plink/>, accessed on 10 May 2023) within R (4.3.0, Posit, PBC, Boston, MA, USA). The QC procedures (see Supplementary Table S1 and Figure S1 for details) can be divided into two parts: sample- and SNP-level filtering. Samples with a genotype missing rate $\geq 2\%$ and missing phenotypes were excluded from the analysis. In addition, SNPs with a genotype missing rate $\geq 2\%$, minor allele frequency (MAF) < 0.01 , and Hardy–Weinberg equilibrium (HWE) exact test p -value $< 10^{-6}$ were filtered out. Finally, 53,224 Taiwanese female participants were included in the following analysis.

2.3. Statistical Analysis

As population genetic structures differ across ethnicities and the Taiwanese population is less representative in international data sources, we randomly split our samples into two datasets for the exploration of genetic association with AAM (trait 1; $N = 26,612$) and T2DM and relevant phenotypes (trait 2; $N = 26,612$). This method is referred to as the Mendelian randomization approach, which is used to minimize confounding effects. We

tested between-group differences to ensure similarities in these two subsets regarding waistline, overall health, body mass index (BMI), BFP, hypertension [16], and years of education [17]. Principal components analysis (PCA) was performed to identify the structure in the distribution of genetic variation across the genetics of all the participants.

Two separate GWASs were conducted for AAM and T2DM traits using two exclusive subsets of data. GWAS for the AAM trait was adjusted for age at enrollment, BMI, and the principal components; GWAS for T2DM traits was adjusted for education level, BMI, and the principal components. Initially, we depicted conditional quantile-quantile (Q-Q) plots to visualize SNP enrichment for T2DM and relevant phenotypes given the different levels of association with AAM. A greater and earlier departure from the reference line (leftward shift) with a higher association with AAM indicates a more significant genetic overlap between AAM and T2DM traits. Further, we applied the conditional false discovery rate (cFDR) method, an extension of the conventional FDR incorporating information from the GWAS summary statistics of a conditional phenotype, to identify the shared genetic variants between AAM and T2DM. In this study, the cFDR is defined as the probability that an SNP is not associated with T2DM traits, given that the p -values for both AAM and T2DM traits are below the significance threshold, and its value indicates the level of pleiotropic enrichment. The cFDR analysis was implemented using an R package (<https://github.com/jamesliley/cFDR-common-controls>, accessed on 4 August 2023) [18]. The shared loci between the AAM and T2DM traits based on the cFDR analysis were visualized in a Manhattan plot. Since multiple significant association p -values in the same region were observed, a linkage disequilibrium (LD)-based clumping procedure was performed to estimate the number of independent loci, with sites less than 500 kb away from an index variant and with r^2 larger than 0.1 assigned to a single clump. To identify the connectivity among the pleiotropic loci, we also used the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <https://string-db.org/>, accessed on 21 August 2023) to build protein–protein interaction (PPI) networks.

3. Results

The average age of the entire study population was 50.9 years. Overall, the average AAM was 13.3 years, and a total of 4.57% reported being diagnosed with T2DM. The two randomly split datasets were similar in most baseline characteristics, age distribution, and risk of T2DM based on the ages of participants (see Supplementary Table S2 and Figure S2 for details).

We depicted conditional Q-Q plots showing SNP enrichment for T2DM traits given different levels of association with AAM (Figure 1). An earlier and greater departure from the reference line was well noted in the plots of BFP and FBG, indicating stronger pleiotropic enrichments being observed for these two traits conditioned on AAM. In contrast, the plots of T2DM diagnosis and HbA_{1c} levels given AAM showed low levels of pleiotropic enrichment. Based on the cFDR analysis, the shared loci between AAM and T2DM and relevant traits were visualized in a Manhattan plot (Figure 2). Using a cutoff of a cFDR less than 0.05, a total of 39 independent shared genetic loci were identified for the four T2DM traits after the clumping procedure (Table 1). The most significant one is *FN3KRP* rs1046896, which has the highest significant level of enrichment (cFDR = 6.84×10^{-49}). In addition, *CDKAL1* rs2206734 (cFDR = 6.48×10^{-10}), *B3GNTL1* rs58431774 (cFDR = 2.95×10^{-10}), *G6PC2* rs1402837 (cFDR = 1.82×10^{-8}), and *KCNQ1* rs60808706 (cFDR = 9.49×10^{-8}) were highlighted for their significant genetic enrichment.

Further, to discover the connectivity among the pleiotropic loci, we built a PPI network using the STRING database, and the network was significantly enriched by showing that the nodes were not random and the observed number of edges was significant (see Supplementary Figure S3 for details). The top 2 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were the insulin signaling pathway and glucagon signaling pathway (see Supplementary Table S3 for details).

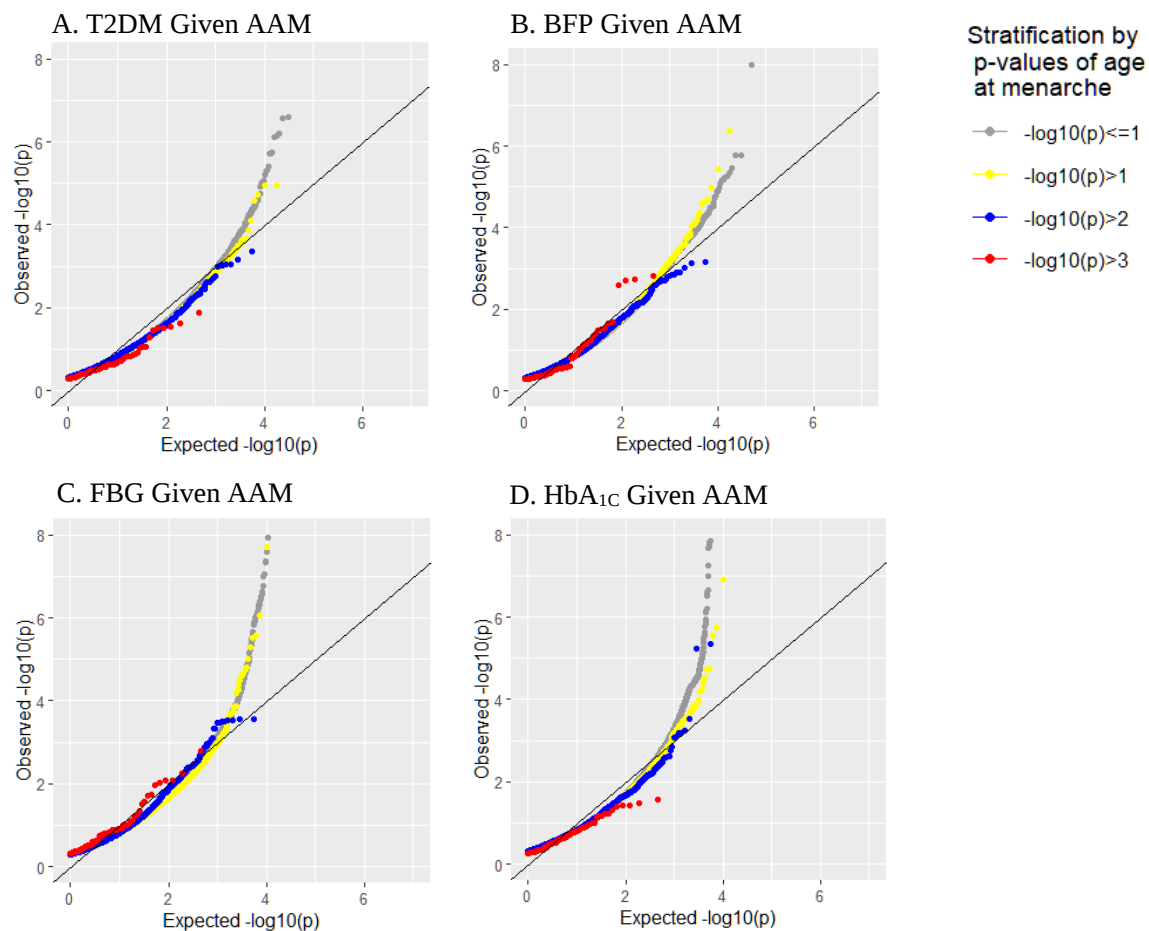


Figure 1. Conditional quantile–quantile plots for diabetic traits conditioned on p -values of AAM. The diagonal line indicates a null hypothesis. (Note: AAM, age at menarche; T2DM, type 2 diabetes mellitus; BFP, body fat percentage; FBG, fasting blood glucose; HbA_{1C}, hemoglobin A1C).

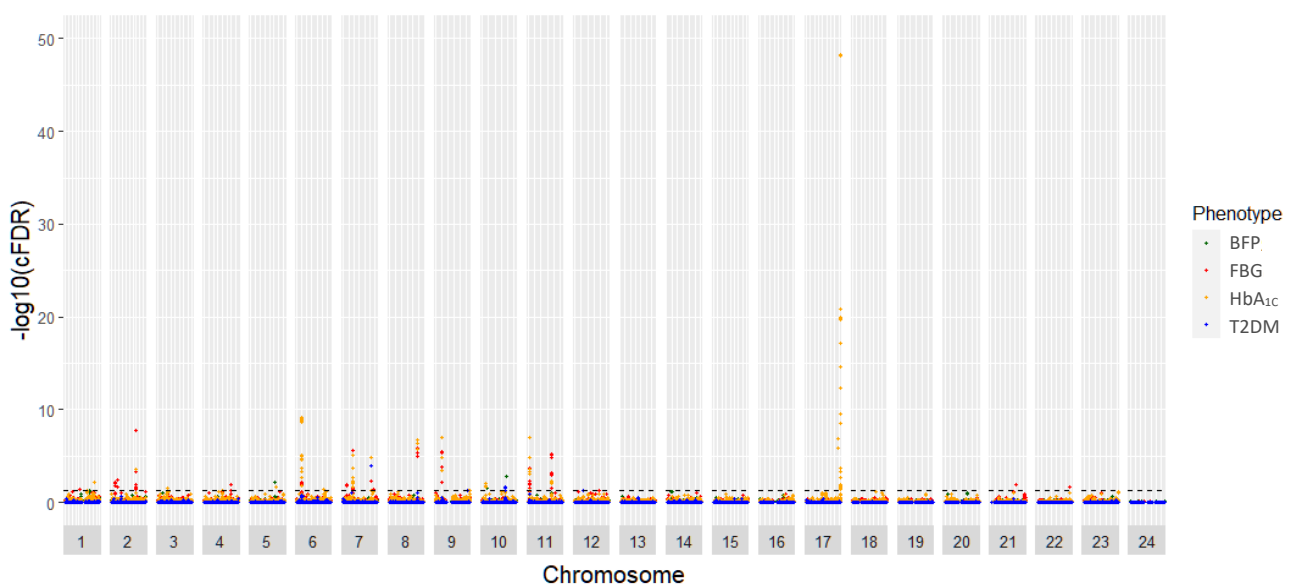


Figure 2. Conditional false discovery rate (cFDR) Manhattan plot of $-\log_{10}$ values for T2DM traits. The dashed line indicates a statistical significance threshold of a cFDR of less than 0.05. (Note: BFP, body fat percentage; FBG, fasting blood glucose; HbA_{1C}, hemoglobin A1C; T2DM, type 2 diabetes mellitus).

Table 1. A total of 39 independent pleiotropic loci in AAM (trait 1) and T2DM (trait 2) with a conditional false discovery rate (cFDR) of less than 0.05. (Note: T2DM, type 2 diabetes mellitus; AAM, age at menarche; Chr, chromosome; hg38, Homo sapiens genome assembly GRCh38; BFP, body fat percentage; FBG, fasting blood glucose; HbA_{1C}, glycated hemoglobin A1C).

Chr	rs	Position (hg38)	Gene	<i>p</i> of Trait 1	Trait 2	<i>p</i> of Trait 2	cFDR
1	rs148500298	1.02×10^8	<i>RP11-202K23.1</i>	0.658659	FBG	5.25×10^{-6}	0.040840737
1	rs3138105	1.58×10^8	<i>CD1B, CD1C</i>	0.475425	HbA _{1C}	9.86×10^{-6}	0.045925272
1	rs17042165	2.15×10^8	<i>LOC124904510</i>	0.00217096	HbA _{1C}	4.61×10^{-6}	0.007599174
2	rs780093	27,519,736	<i>GCKR</i>	0.320034	FBG	4.80×10^{-7}	0.007515765
2	rs895636	44,961,214	<i>AC012354.6</i>	0.807344	FBG	1.67×10^{-7}	0.003508871
2	rs1402837	1.69×10^8	<i>G6PC2</i>	0.280055	HbA _{1C}	1.44×10^{-8}	0.00028397
2	rs116933981	1.69×10^8	<i>LRP2</i>	0.41804	FBG	1.32×10^{-13}	1.82×10^{-8}
2	rs116933981	1.69×10^8	<i>LRP2</i>	0.41804	FBG	8.63×10^{-7}	0.012442801
3	rs3804766	51,393,226	<i>RBM15B</i>	0.208714	HbA _{1C}	3.20×10^{-6}	0.033487115
4	rs223461	1.03×10^8	<i>LOC102723704, LOC124900743</i>	0.192597	HbA _{1C}	5.45×10^{-6}	0.048127897
4	rs79196252	1.47×10^8	<i>SLC10A7</i>	0.0439132	FBG	8.66×10^{-7}	0.010702327
5	rs248062	1.27×10^8	<i>CTB-1121.1</i>	0.0326756	BFP	4.10×10^{-7}	0.007739594
5	rs75170429	1.33×10^8	<i>FSTL4</i>	0.419274	HbA _{1C}	2.51×10^{-6}	0.019516832
6	rs2206734	20,694,653	<i>CDKAL1</i>	0.712727	HbA _{1C}	1.94×10^{-14}	6.49×10^{-10}
6	rs9376090	1.35×10^8	<i>HBS1L</i>	0.695327	FBG	3.05×10^{-7}	0.006010452
6	rs9376090	1.35×10^8	<i>HBS1L</i>	0.695327	HbA _{1C}	7.17×10^{-6}	0.041947507
7	rs10244051	15,024,208	<i>GTF3AP5</i>	0.776841	FBG	8.57×10^{-7}	0.012993208
7	rs16881016	44,171,808	<i>GCK, LOC105375257</i>	0.480132	HbA _{1C}	4.52×10^{-10}	8.01×10^{-6}
7	rs1799884	44,189,469	<i>GCK</i>	0.264629	FBG	2.01×10^{-11}	2.31×10^{-6}
7	rs6975024	44,192,287	<i>GCK</i>	0.601593	FBG	3.12×10^{-6}	0.029224933
7	rs6975024	44,192,287	<i>GCK</i>	0.601593	HbA _{1C}	2.22×10^{-7}	0.002111339
7	rs2233580	1.28×10^8	<i>PAX4</i>	0.403983	T2DM	5.07×10^{-10}	9.90×10^{-5}
7	rs2233580	1.28×10^8	<i>PAX4</i>	0.403983	HbA _{1C}	6.18×10^{-10}	1.51×10^{-5}
7	rs6943771	1.4×10^8	<i>PARP12</i>	0.995326	FBG	2.70×10^{-7}	0.005270602
7	rs6943771	1.4×10^8	<i>PARP12</i>	0.995326	FBG	4.37×10^{-6}	0.038811336
8	rs3802177	1.17×10^8	<i>SLC30A8, LOC105375716</i>	0.889634	HbA _{1C}	8.07×10^{-12}	1.67×10^{-7}
8	rs3802177	1.17×10^8	<i>SLC30A8, LOC105375716</i>	0.889634	FBG	7.81×10^{-12}	1.62×10^{-6}
9	rs10965250	22,133,285	<i>CDKN2B-AS1</i>	0.236788	HbA _{1C}	8.65×10^{-13}	1.02×10^{-7}
9	rs10965250	22,133,285	<i>CDKN2B-AS1</i>	0.236788	FBG	3.00×10^{-11}	3.54×10^{-6}
9	rs1050700	1.33×10^8	<i>TSC1</i>	0.487359	HbA _{1C}	1.16×10^{-5}	0.049462307
10	rs12221133	12,211,598	<i>CDC123</i>	0.45323	HbA _{1C}	1.44×10^{-6}	0.010460355
10	rs151268010	16,675,558	<i>RSU1</i>	0.0122221	BFP	3.57×10^{-6}	0.027178562
10	rs75631171	92,449,649	<i>IDE</i>	0.190513	T2DM	2.55×10^{-7}	0.024501519
10	rs1111875	92,703,125	<i>Y_RNA</i>	0.664192	T2DM	2.59×10^{-7}	0.027071426
10	rs10786156	94,254,865	<i>PLCE1</i>	0.332509	BFP	1.02×10^{-8}	0.00165048
11	rs60808706	2,836,003	<i>KCNQ1</i>	0.589924	HbA _{1C}	3.38×10^{-12}	9.49×10^{-8}
11	rs10466351	92,964,815	<i>LOC124902733</i>	0.44799	FBG	6.33×10^{-9}	0.000221717
11	rs10830963	92,975,544	<i>MTNR1B</i>	0.526219	HbA _{1C}	1.25×10^{-10}	6.76×10^{-6}
11	rs10830963	92,975,544	<i>MTNR1B</i>	0.526219	HbA _{1C}	6.32×10^{-7}	0.005298336
17	rs761772	78,125,997	<i>TMC6</i>	0.429945	HbA _{1C}	4.79×10^{-12}	1.42×10^{-7}
17	rs73357173	82,670,096	<i>RAB40B/MIR4525</i>	0.0342794	HbA _{1C}	2.89×10^{-6}	0.019044672
17	rs2250754	82,703,440	<i>LOC124904093</i>	0.0102342	HbA _{1C}	1.73×10^{-6}	0.01124699
17	rs1046896	82,727,657	<i>FN3KRP</i>	0.382707	HbA _{1C}	3.68×10^{-54}	6.84×10^{-49}
17	rs3785519	82,937,832	<i>TBCD</i>	0.177841	HbA _{1C}	2.08×10^{-8}	0.000467929
17	rs58431774	83,000,587	<i>B3GNTL1</i>	0.742557	HbA _{1C}	7.61×10^{-15}	2.95×10^{-10}
21	rs7510550	36,629,993	<i>AP000696.2</i>	0.515658	FBG	7.92×10^{-7}	0.01085386
22	rs146847831	48,773,800	<i>FAM19A5</i>	0.106661	FBG	1.42×10^{-6}	0.019983018

4. Discussion

In this study, a genome-wide statistical approach was used to identify 39 shared loci between AAM and T2DM traits, with some reported in association with T2DM [19–23].

The conditional Q-Q plots showed pleiotropic enrichments for BFR and FBG conditioned on AAM, but the enrichment effect was less prominent for T2DM diagnosis and HbA_{1C} levels. This discrepancy may result from data sources and availability. The diagnosis was self-reported, and the medication history was unavailable in TWB. There was a possibility that T2DM was undiagnosed or diabetic types were misclassified by participants. Moreover, HbA_{1C} might reflect glycemic control corresponding to patients' medication adherence rather than disease severity among patients with T2DM. Obtaining collateral reports from treating physicians or linked medication records may ameliorate data quality and accuracy.

Using the cFDR analysis, our study confirmed some novel enrichment loci in shared genetics between AAM and T2DM traits, such as rs1046896 (*FN3KRP*), rs2206734 (*CDKAL1*), rs58431774 (*B3GNTL1*), rs1402837 (*G6PC2*), and rs60808706 (*KCNQ1*). The SNP rs1046896 (*FN3KRP*) has the highest genetic enrichment level. *FN3KRP* (fructosamine-3-kinase-related protein) is involved in the deglycation of proteins modified via non-enzymatic glycation [24]. A high glucose concentration can result in the non-enzymatic oxidation of proteins via the reaction of glucose-6-phosphate and lysine residues. Proteins modified in this way become less active or functional. This reaction is referred to as the non-enzymatic glycation of proteins or the Maillard reaction. Hyperglycemia is a primary factor that promotes glycation, and the glycation end product forms on DNA, lipids, and proteins, where they represent pathophysiological modifications that precipitate dysfunction at a cellular and molecular level [25]. This further leads to several complications, such as blindness, heart disease, nerve damage, and kidney failure [26]. *FN3KRP* may result in the deglycation of proteins to restore their function.

CDKAL1 (cyclin-dependent kinase 5 regulatory subunit associated protein 1 like 1) has been reported as a major pathogenesis-related protein for T2DM [27]. The function of *CDKAL1* has not been entirely determined. Research has shown that *CDKAL1* may be involved in beta cell dysfunction in the pancreas and the regulation of mitochondrial function in adipose tissue [28,29].

B3GNTL1 (beta1,3-N-Acetylglucosaminyltransferase-Like Protein 1) protein is predicted to enable glycosyltransferase activity [30]. According to a recent study, analysis of the PPI network of the genes containing *B3GNTL1* reveals that, at the molecular level, there seemed to be interconnected factors that affect the progression of renal impairment among diabetic patients [31].

G6PC2 (glucose-6-phosphatase catalytic subunit 2) catalyzes the hydrolysis of glucose-6-phosphate to produce glucose and inorganic phosphate. The gene is explicitly expressed in pancreatic islet beta cells. The mutation in *G6PC2* was found to have a lead role in the modulation of FBG levels, and, thus, it might increase the risk of T2DM [32]. In addition, previous GWAS showed linked polymorphisms in *G6PC2*, with variations in FBG and BFP [33].

KCNQ1 was previously identified for its correlation with T2DM [34]. *KCNQ1* encodes a voltage-gated potassium channel required for the repolarization phase of the action potential of cardiac muscles [35], and the gene is also expressed in other tissues, including the brain, adipose tissue, and pancreas, as well as in the insulin-secreting cell line [36]. In addition to its well-studied association with T2DM, research has revealed a significant association with menopausal age [37].

The PPI analysis revealed that the whole network was significantly enriched. The proteins at the center of the PPI network were insulin (encoded by *INS*) and AKT serine/threonine kinase 1 (encoded by *AKT1*). Both are involved in glucose homeostasis and metabolism [38]. Moreover, the top two KEGG pathways were the insulin and glucagon signaling pathways. Insulin plays a critical role in multiple physiological processes, such as promoting glucose uptake, glycogen synthesis, lipogenesis, and protein synthesis. Reduced production of insulin and insulin resistance are the main causes of diabetes. Insulin is an anabolic peptide hormone secreted by pancreatic β cells acting through a receptor located in the membrane of target cells. The receptor activates a complex intracellular signaling network. The two main pathways of insulin signaling emanating from the

insulin receptor-IRS node are the PI3K (phosphatidylinositol 3-kinase)/AKT (AKT serine/threonine kinase 1) pathway and the Raf/Ras/MEK (mitogen-activated protein kinase kinase)/MAPK (mitogen-activated protein kinase) pathway [39]. Glucagon is also an important regulator of glucose homeostasis, and it is secreted by pancreatic α cells. Once glucagon binds to a transmembrane receptor on target cells, it leads to adenylate cyclase activation and cAMP formation. The increase in intracellular cAMP levels activates PKA (protein kinase A), which phosphorylates CREB (cAMP response element binding protein), PFK-2 (phospho-fructokinase 2)/FBPase2 (fructose 2,6-bisphosphatase), pyruvate kinase, phosphorylase kinase, and glycogen synthase. Overall, glucagon increases gluconeogenesis and glycogenolysis and decreases glycolysis [40]. Most of the novel-discovered genes were also found to be involved in these pathways. For example, *CREB* encodes the protein that regulates the gene expression of the enzymes of gluconeogenesis, and it has recently been regarded as a targeting factor in diabetic treatment [41].

Our findings may have implications for clinical practices and directions for future research. Firstly, genetic screening for the shared loci between AAM and T2DM may inform lifestyle modification and glycemic monitoring among early-maturing women. In this sense, further exploration of the gene–lifestyle interaction effects on T2DM may help to identify high-risk groups targeted for advanced pharmacotherapy. Moreover, enriched pathways involving shared genetics may shed light on experimental research into T2DM mechanisms associated with early AAM, a critical step toward personalized diabetic prevention and medication strategies.

Despite its population size and representativeness strength, this study has several limitations. Firstly, the genetic association analysis was derived from and, thus, may be limited to the Taiwanese population. Cross-ethnic validation may be needed, as genetic associations with AAM differ between Asian and European women [42]. Secondly, AAM and diagnosis of T2DM in TWB were based on self-reported questionnaires. Recall bias may exist. Thirdly, the use of hypoglycemic medications was not available in TWB. An augmented dataset linking the TWB and insurance claims of the National Health Insurance Research Database may help to clarify these issues. Fourthly, the GWAS results could be affected by selection bias, as individuals who volunteer to participate may not be representative of the underlying sampling population.

5. Conclusions

This study highlighted potential pleiotropic effects across AAM and T2DM that may explain the association between the two phenotypes. The result may suggest genetic screening for the enriched loci that may help to inform risks for T2DM in women with early AAM. The shared genetic variants discovered are targets for the future exploration of early AAM and T2DM mechanisms.

Supplementary Materials: The following supporting information can be downloaded via this link: <https://www.mdpi.com/article/10.3390/biomedicines12010157/s1>, Figure S1: Description of quality control procedures; Figure S2: Risk of T2DM based on age of participants; Figure S3: Protein-protein interaction (PPI) network for novel pleiotropic loci; Table S1: Description of quality control procedures; Table S2: General characteristics of participants in 2 datasets; Table S3: Top 10 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways.

Author Contributions: M.-C.T. conceived the study. Y.-F.C. conducted, C.-Y.Y. reviewed, and M.-C.T. supervised the statistical analysis. Y.-F.C. drafted and M.-C.T. critically reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the research grant awarded to Tsai MC by the National Cheng Kung University Hospital (NCKUH-10902012) and the summer research project grant awarded to Cheng YF by the College of Medicine at National Cheng Kung University (NCKUMCS2023075).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of National Cheng Kung University (A-ER-108-331).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The TWB dataset is available upon application for research purposes. A detailed description of TWB data availability and the application process could be found via the following link: https://taiwanview.twbiobank.org.tw/data_appl.

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. International Diabetes Federation—Home. idf.org. Available online: <https://idf.org/> (accessed on 25 March 2023).
2. Health Promotion Administration, Taiwan. 31 December 2016. Available online: <https://www.hpa.gov.tw/Pages/List.aspx?nodeid=359> (accessed on 25 March 2023).
3. Hawkes, C.H. Twin Studies in Diabetes Mellitus. *Diabet. Med.* **1997**, *14*, 347–352. [CrossRef]
4. Ali, O. Genetics of type 2 diabetes. *World J. Diabetes* **2013**, *4*, 114–123. [CrossRef] [PubMed]
5. Chow, J.C.; Chou, T.Y.; Tung, T.-H.; Yuh, Y.-S. Recent pubertal timing trends in Northern Taiwanese children: Comparison with skeletal maturity. *J. Chin. Med. Assoc.* **2020**, *83*, 870–875. [CrossRef] [PubMed]
6. Lee, H.S. Why should we be concerned about early menarche? *Clin. Exp. Pediatr.* **2021**, *64*, 26–27. [CrossRef]
7. Karapanou, O.; Papadimitriou, A. Determinants of menarche. *Reprod. Biol. Endocrinol.* **2010**, *8*, 115. [CrossRef] [PubMed]
8. Shadyab, A.H.; Macera, C.A.; Shaffer, R.A.; Jain, S.; Gallo, L.C.; Gass, M.L.; Waring, M.E.; Stefanick, M.L.; LaCroix, A.Z. Ages at menarche and menopause and reproductive lifespan as predictors of exceptional longevity in women: The Women’s Health Initiative. *Menopause* **2017**, *24*, 35–44. [CrossRef]
9. CDC. *Learn How People with Type 2 Diabetes Can Live Longer*; Centers for Disease Control and Prevention: Atlanta, GA, USA, 2022. Available online: <https://www.cdc.gov/diabetes/resources-publications/research-summaries/reaching-treatment-goals.html> (accessed on 25 March 2023).
10. Elks, C.E.; Ong, K.K.; Scott, R.A.; van der Schouw, Y.T.; Brand, J.S.; Wark, P.A.; Amiano, P.; Balkau, B.; Barricarte, A.; Boeing, H.; et al. Age at Menarche and Type 2 Diabetes Risk. *Diabetes Care* **2013**, *36*, 3526–3534. [CrossRef]
11. Feng, Y.-C.A.; Chen, C.-Y.; Chen, T.-T.; Kuo, P.-H.; Hsu, Y.-H.; Yang, H.-I.; Chen, W.J.; Su, M.-W.; Chu, H.-W.; Shen, C.-Y.; et al. Taiwan Biobank: A rich biomedical research database of the Taiwanese population. *Cell Genom.* **2022**, *2*, 100197. [CrossRef]
12. Fan, C.-T.; Hung, T.-H.; Yeh, C.-K. Taiwan Regulation of Biobanks. *J. Law Med. Ethics* **2015**, *43*, 816–826. [CrossRef]
13. Liao, S.-F.; Su, C.-Y.; Su, M.-H.; Chen, C.-Y.; Chen, C.-Y.; Lin, Y.-F.; Pan, Y.-J.; Hsiao, P.-C.; Chen, P.-C.; Huang, Y.-T.; et al. Associations of polygenic risks, depression, and obesity-related traits in Taiwan Biobank. *J. Affect. Disord.* **2023**, *320*, 397–403. [CrossRef]
14. Tsai, M.-C.; Hsu, C.-H.; Chu, S.-K.; Roy-Gagnon, M.-H.; Lin, S.-H. Genome-wide association study of age at menarche in the Taiwan Biobank suggests NOL4 as a novel associated gene. *J. Hum. Genet.* **2023**, *68*, 339–345. [CrossRef] [PubMed]
15. Lin, W.-Y.; Chan, C.-C.; Liu, Y.-L.; Yang, A.C.; Tsai, S.-J.; Kuo, P.-H. Performing different kinds of physical exercise differentially attenuates the genetic effects on obesity measures: Evidence from 18,424 Taiwan Biobank participants. *PLoS Genet.* **2019**, *15*, e1008277. [CrossRef] [PubMed]
16. Vujkovic, M.; Keaton, J.M.; Lynch, J.A.; Miller, D.R.; Zhou, J.; Tcheandjieu, C.; Huffman, J.E.; Assimes, T.L.; Lorenz, K.; Zhu, X.; et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat. Genet.* **2020**, *52*, 680–691. [CrossRef] [PubMed]
17. Borrell, L.N.; Dallo, F.J.; White, K. Education and Diabetes in a Racially and Ethnically Diverse Population. *Am. J. Public Health* **2006**, *96*, 1637–1642. [CrossRef] [PubMed]
18. Liley, J.; Wallace, C. A Pleiotropy-Informed Bayesian False Discovery Rate Adapted to a Shared Control Design Finds New Disease Associations from GWAS Summary Statistics. *PLoS Genet.* **2015**, *11*, e1004926. [CrossRef]
19. Basile, K.J.; Johnson, M.E.; Xia, Q.; Grant, S.F.A. Genetic Susceptibility to Type 2 Diabetes and Obesity: Follow-Up of Findings from Genome-Wide Association Studies. *Int. J. Endocrinol.* **2014**, *2014*, 769671. [CrossRef] [PubMed]
20. Sanghera, D.K.; Blackett, P.R. Type 2 Diabetes Genetics: Beyond GWAS. *J. Diabetes Metab.* **2012**, *3*, 6948. [CrossRef]
21. Burillo, J.; Marqués, P.; Jiménez, B.; González-Blanco, C.; Benito, M.; Guillén, C. Insulin Resistance and Diabetes Mellitus in Alzheimer’s Disease. *Cells* **2021**, *10*, 1236. [CrossRef]
22. Demir, S.; Nawroth, P.P.; Herzig, S.; Üstünel, B.E. Emerging Targets in Type 2 Diabetes and Diabetic Complications. *Adv. Sci.* **2021**, *8*, 2100275. [CrossRef]
23. Chen, C.; Han, M.; Zhang, W.; Cui, J.; Cui, N.; Cao, L.; Luo, G.; Sun, J. Identification of key genes and pathways in type 2 diabetes mellitus and vitamin C metabolism through bioinformatics analysis. *Asia Pac. J. Clin. Nutr.* **2021**, *30*, 715–729. [CrossRef] [PubMed]
24. FN3KRP Fructosamine 3 Kinase Related Protein [Homo Sapiens (Human)]—Gene—NCBI. Available online: <https://www.ncbi.nlm.nih.gov/gene/79672#summary> (accessed on 22 November 2023).
25. Szwergold, B.S. Fructosamine-6-phosphates are deglycated by phosphorylation to fructosamine-3,6-bisphosphates catalyzed by fructosamine-3-kinase (FN3K) and/or fructosamine-3-kinase-related-protein (FN3KRP). *Med. Hypotheses* **2006**, *68*, 37–45. [CrossRef] [PubMed]

26. Ansari, N.A.; Rasheed, Z. Non-enzymatic glycation of proteins: From diabetes to cancer. *Biochem. (Moscow) Suppl. Ser. B Biomed. Chem.* **2009**, *3*, 335–342. [\[CrossRef\]](#)
27. Zeng, Q.; Zou, D.; Gu, S.; Han, F.; Cao, S.; Wei, Y.; Guo, R. Different Associations Between CDKAL1 Variants and Type 2 Diabetes Mellitus Susceptibility: A Meta-analysis. *Front. Genet.* **2022**, *12*, 783078. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Xu, N.; Zhang, T.-T.; Han, W.-J.; Yin, L.-P.; Ma, N.-Z.; Shi, X.-Y.; Sun, J.-J. Association of CDKAL1 RS10946398 Gene Polymorphism with Susceptibility to Diabetes Mellitus Type 2: A Meta-Analysis. *J. Diabetes Res.* **2021**, *2021*, 1254968. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Palmer, C.J.; Bruckner, R.J.; Paulo, J.A.; Kazak, L.; Long, J.Z.; Mina, A.I.; Deng, Z.; LeClair, K.B.; Hall, J.A.; Hong, S.; et al. Cdkal1, a type 2 diabetes susceptibility gene, regulates mitochondrial function in adipose tissue. *Mol. Metab.* **2017**, *6*, 1212–1225. [\[CrossRef\]](#) [\[PubMed\]](#)
30. B3GNTL1 UDP-GlcNAc:betaGal Beta-1,3-N-acetylglucosaminyltransferase Like 1 [Homo sapiens (human)]—Gene—NCBI. Available online: <https://www.ncbi.nlm.nih.gov/gene/146712> (accessed on 22 November 2023).
31. Wu, I.-W.; Tsai, T.-H.; Lo, C.-J.; Chou, Y.-J.; Yeh, C.-H.; Chan, Y.-H.; Chen, J.-H.; Hsu, P.W.-C.; Pan, H.-C.; Hsu, H.-J.; et al. Discovering a trans-omics biomarker signature that predisposes high risk diabetic patients to diabetic kidney disease. *npj Digit. Med.* **2022**, *5*, 166. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Udhaya Kumar, S.; Kamaraj, B.; Varghese, R.P.; Preethi, V.A.; Bithia, R.; George Priya Doss, C. Mutations in G6PC2 gene with increased risk for development of type 2 diabetes: Understanding via computational approach. *Adv. Protein Chem. Struct. Biol.* **2022**, *130*, 351–373. [\[CrossRef\]](#)
33. Bouatia-Naji, N.; Rocheleau, G.; Van Lommel, L.; Lemaire, K.; Schuit, F.; Cavalcanti-Proença, C.; Marchand, M.; Hartikainen, A.-L.; Sovio, U.; De Graeve, F.; et al. A Polymorphism within the G6PC2 Gene Is Associated with Fasting Plasma Glucose Levels. *Science* **2008**, *320*, 1085–1088. [\[CrossRef\]](#)
34. Kasuga, M. KCNQ1, a susceptibility gene for type 2 diabetes. *J. Diabetes Investig.* **2011**, *2*, 413–414. [\[CrossRef\]](#)
35. KCNQ1 Potassium Voltage-Gated Channel Subfamily Q Member 1 [Homo Sapiens (Human)]—Gene—NCBI. Available online: <https://www.ncbi.nlm.nih.gov/gene/3784> (accessed on 7 October 2023).
36. Zhang, W.; Wang, H.; Guan, X.; Niu, Q.; Li, W. Variant rs2237892 of KCNQ1 Is Potentially Associated with Hypertension and Macrovascular Complications in Type 2 Diabetes Mellitus in A Chinese Han Population. *Genom. Proteom. Bioinform.* **2015**, *13*, 364–370. [\[CrossRef\]](#)
37. Hardy, T.M.; de Mendoza, V.B.; Sun, Y.V.; Taylor, J.Y. Genomics of Reproductive Traits and Cardiometabolic Disease Risk in African American Women. *Nurs. Res.* **2019**, *68*, 135–144. [\[CrossRef\]](#)
38. Hay, N. Akt isoforms and glucose homeostasis—The leptin connection. *Trends Endocrinol. Metab.* **2011**, *22*, 66–73. [\[CrossRef\]](#)
39. De Meyts, P.; Feingold, K.R.; Anawalt, B.; Boyce, A.; Chrousos, G.; de Herder, W.W.; Dhatariya, K.; Dungan, K.; Hershman, J.M.; Hofland, J.; et al. (Eds.) The Insulin Receptor and Its Signal Transduction Network. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK378978/> (accessed on 6 March 2022).
40. Janah, L.; Kjeldsen, S.; Galsgaard, K.D.; Winther-Sørensen, M.; Stojanovska, E.; Pedersen, J.; Knop, F.K.; Holst, J.J.; Albrechtsen, N.J.W. Glucagon Receptor Signaling and Glucagon Resistance. *Int. J. Mol. Sci.* **2019**, *20*, 3314. [\[CrossRef\]](#)
41. Benchoula, K.; Parhar, I.S.; Madhavan, P.; Hwa, W.E. CREB nuclear transcription activity as a targeting factor in the treatment of diabetes and diabetes complications. *Biochem. Pharmacol.* **2021**, *188*, 114531. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Yermachenko, A.; Dvornyk, V. Nongenetic Determinants of Age at Menarche: A Systematic Review. *BioMed Res. Int.* **2014**, *2014*, 371583. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.