Genetically predicted 1091 blood metabolites and 309 metabolite ratios in relation to risk of type 2 diabetes: a Mendelian randomization study

Jixin Li[1†], Wenru Wang[1†], Fengzhao Liu[2†], Linjie Qiu[1], Yan Ren[1], Meijie Li[1], Wenjie Li[1], Feng Gao[1]* and Jin Zhang[1]*

[1]Xiyuan Hospital of the China Academy of Chinese Medical Sciences, Beijing, China, [2]First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, China

Background: Metabolic dysregulation represents a defining characteristic of Type 2 diabetes (T2DM). Nevertheless, there remains an absence of substantial evidence establishing a direct causal link between circulating blood metabolites and the promotion or prevention of T2DM. In addressing this gap, we employed Mendelian randomization (MR) analysis to investigate the potential causal association between 1,091 blood metabolites, 309 metabolite ratios, and the occurrence of T2DM.

Methods: Data encompassing single-nucleotide polymorphisms (SNPs) for 1,091 blood metabolites and 309 metabolite ratios were extracted from a Canadian Genome-wide association study (GWAS) involving 8,299 participants. To evaluate the causal link between these metabolites and Type 2 diabetes (T2DM), multiple methods including Inverse Variance Weighted (IVW), Weighted Median, MR Egger, Weighted Mode, and Simple Mode were employed. p-values underwent correction utilizing False Discovery Rates (FDR). Sensitivity analyses incorporated Cochran's Q test, MR-Egger intercept test, MR-PRESSO, Steiger test, leave-one-out analysis, and single SNP analysis. The causal effects were visualized via Circos plot, forest plot, and scatter plot. Furthermore, for noteworthy, an independent T2DM GWAS dataset (GCST006867) was utilized for replication analysis. Metabolic pathway analysis of closely correlated metabolites was conducted using MetaboAnalyst 5.0.

Results: The IVW analysis method utilized in this study revealed 88 blood metabolites and 37 metabolite ratios demonstrating a significant causal relationship with T2DM ($p < 0.05$). Notably, strong causal associations with T2DM were observed for specific metabolites: 1-linoleoyl-GPE (18:2) (IVW: OR:0.930, 95% CI: 0.899−0.962, $p = 2.16 \times 10^{-5}$), 1,2-dilinoleoyl-GPE (18:2/18:2) (IVW: OR:0.942, 95% CI: 0.917−0.968, $p = 1.64 \times 10^{-5}$), Mannose (IVW: OR: 1.133, 95% CI: 1.072−1.197, $p = 1.02 \times 10^{-5}$), X-21829 (IVW: OR:1.036, 95% CI: 1.036−1.122, $p = 9.44 \times 10^{-5}$), and Phosphate to mannose ratio (IVW: OR:0.870, 95% CI: 0.818−0.926, $p = 1.29 \times 10^{-5}$, FDR = 0.008). Additionally, metabolic pathway analysis highlighted six significant pathways associated with T2DM development: Valine, leucine and isoleucine biosynthesis, Phenylalanine

metabolism, Glycerophospholipid metabolism, Alpha-Linolenic acid metabolism, Sphingolipid metabolism, and Alanine, aspartate, and glutamate metabolism.

Conclusion: This study identifies both protective and risk-associated metabolites that play a causal role in the development of T2DM. By integrating genomics and metabolomics, it presents novel insights into the pathogenesis of T2DM. These findings hold potential implications for early screening, preventive measures, and treatment strategies for T2DM.

1 Introduction

New data released by the International Diabetes Federation (IDF) in 2021 indicates a global prevalence of 537 million adults diagnosed with diabetes mellitus (DM). This represents a 16% increase from the 2019 forecast. Projections suggest a further rise to 643 million by 2030 and a staggering 783 million by 2045 (Ogurtsova et al., 2022). Type 2 diabetes (T2DM) manifests as a metabolic disorder typified by both insulin resistance and insufficient insulin secretion (Keller-Pinter et al., 2023). Insulin resistance represents an initial stage in the progression of T2DM. It affects approximately 40% of young adults in the United States (Parcha et al., 2022). The World Health Organization predicts that the global number of T2DM patients will double, reaching 350 million by 2030 (Collins et al., 2011). Type 2 diabetes mellitus (T2DM) is associated with complications like renal disease, coronary heart disease, peripheral vascular disease, and other systemic issues, causing a mortality rate twice as high as that of the healthy population as the disease progresses (Mulnier et al., 2006). In recent years, T2DM has shown a trend towards affecting individuals at a younger age, exhibiting prolonged disease duration, multiple complications, and significant risks (Magliano et al., 2020). It imposes a substantial economic burden globally and stands as a critical public health concern. Identifying risk factors for T2DM and proactively managing high-risk individuals holds immense significance in preventing and treating this condition.

Insulin resistance stands prominently as a pivotal factor in the pathogenesis of T2DM, driven not solely by the compromised transmission of the insulin signaling pathway but also entailing a myriad of intricate metabolic determinants (Yang et al., 2018). For instance, saturated fatty acids exhibit the potential to impair insulin sensitivity in murine models (Holland et al., 2011). Conversely, judicious intake of unsaturated fatty acids holds promise in ameliorating insulin sensitivity and mitigating the risk of T2DM (Imamura et al., 2016). Moreover, more than 4 decades ago, it was discerned that heightened serum levels of certain amino acids strongly correlate with obesity and insulin resistance (Felig et al., 1969). Metabolites serve as the foundation of biological expression, providing crucial insights into metabolic processes and disease pathogenesis (Bauermeister et al., 2022; Johnson et al., 2016). Metabolomics allows for the exploration of metabolic pathways or networks through qualitative and quantitative analysis of metabolites, unveiling the metabolic and reaction mechanisms of various diseases, drugs, and chemicals across different organisms (Johnson et al., 2016). Metabolomics enables a thorough characterization of serum metabolite

alterations in patients with T2DM both pre- and post-onset, as well as following treatment initiation. Lee (Newgard et al., 2009), through untargeted metabolomics, identified the plasma metabolite Branched Chain Amino Acid (BCAA), showing a negative correlation with insulin sensitivity and insulin metabolic clearance. Similarly, Marina (Mora-Ortiz et al., 2022) identified 12 metabolites predictive of T2DM remission over a 5-year dietary intervention in 190 T2DM patients. This highlights how metabolomics technology facilitates understanding the mechanisms underlying diabetes development. While numerous studies have explored the correlation between metabolites and T2DM(Bloomgarden, 2018; Long et al., 2020; Mora-Ortiz et al., 2022; Newgard et al., 2009; Palmer et al., 2015; Curtin and Schulz, 1998; Wang et al., 2017), the coverage of metabolites remains incomplete. None have comprehensively accounted for the influence of confounding factors, potentially leading to biased results.

Genomics and metabolomics are closely intertwined. Their integration enables the identification of key mechanisms underlying the development of T2DM and facilitates the exploration of disease markers and drug targets. Mendelian randomization (MR) serves as an epidemiological and genetically grounded method, exploring causal links between exposures and outcomes. Adhering to Mendel's second law, MR studies are not influenced by confounding factors due to the random assignment of alleles during gamete formation. This study integrates genomics and metabolomics, leveraging the latest comprehensive genome-wide association study (GWAS) on blood metabolites (Chen et al., 2023). It employs Mendelian randomization (MR) analysis to investigate the causal associations of 1,091 blood metabolites and 309 metabolite ratios with T2DM. The objective is to comprehensively dissect the pathogenesis of T2DM and its metabolic pathways, thereby providing insights into prediction, diagnosis, and treatment.
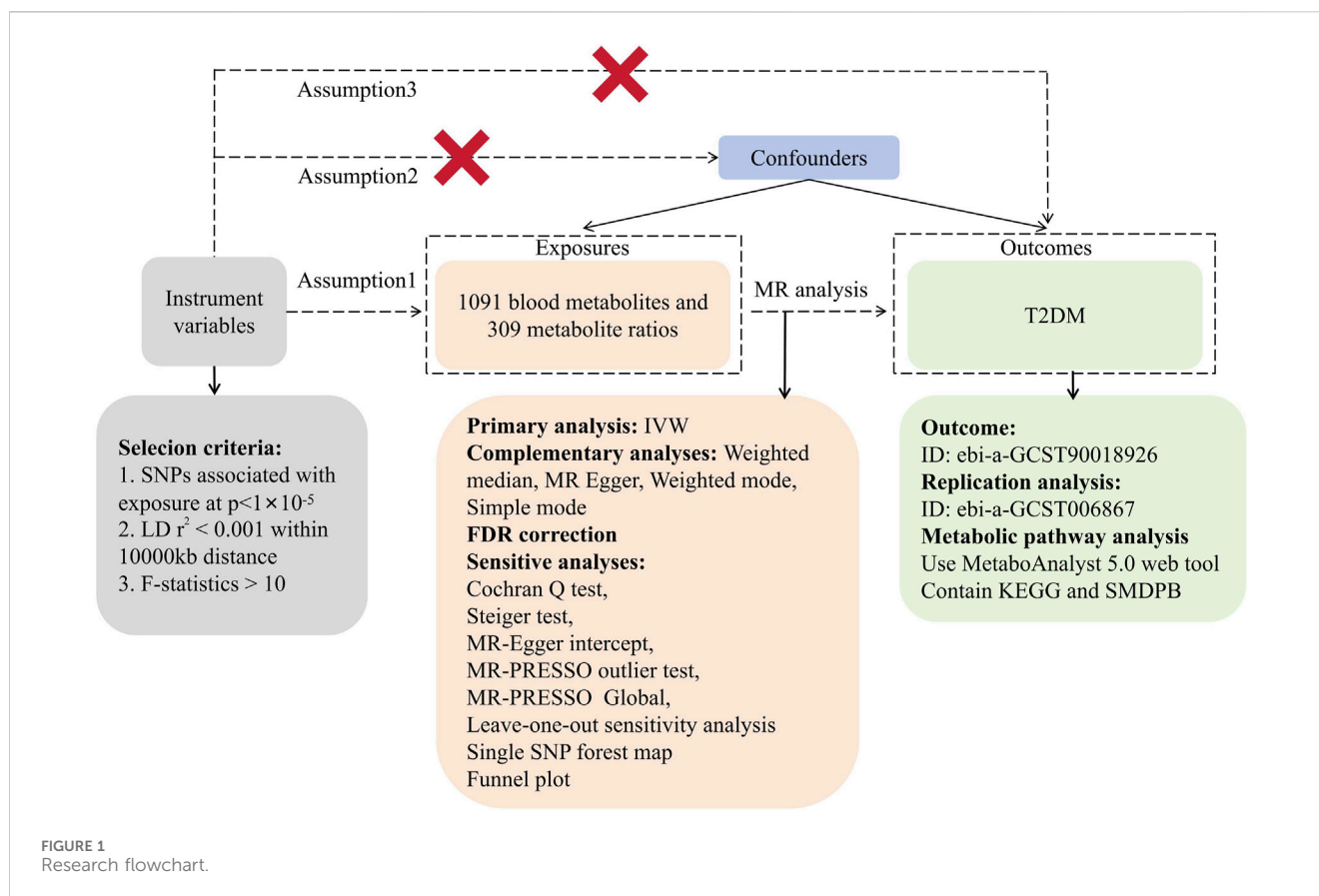
2 Materials and methods

2.1 Reporting guidelines

This study was reported in strict adherence to STROBE-MR guidelines (Skrivankova et al., 2021).

2.2 Research design

In this investigation, we utilized 1,091 blood metabolites and 309 metabolite ratios as 'exposures' while considering T2DM as the

FIGURE 1
Research flowchart.

outcome. Instrumental variables (IVs) were meticulously screened for MR analysis. The study's consistency underwent evaluation using the Cochran Q test. Sensitivity analyses encompassed horizontal multiplicity analysis and a 'leave-one-out' approach, reinforcing the reliability of our findings. Mendelian randomization (MR) studies require adherence to three fundamental assumptions: (1) establishing a robust correlation between IVs and exposure, (2) ensuring IVs' independence from any confounding factors associated with exposure and outcome, and (3) affirming that IVs solely impact outcomes through exposure pathways. In our investigation, MR analysis was employed to ascertain the causal relationship between 1,091 blood metabolites, 309 metabolite ratios, and T2DM. Figure 1 illustrates the study's workflow (Figure 1).

2.3 Data sources

The instrumental variables for this study were sourced from Chen's Canadian Longitudinal Study of Aging (CLSA), encompassing a sample size of 8,299 (Chen et al., 2023). CLSA recruited 51,338 Canadians aged 45–85 years, incorporating concurrent statistical data on their physiology, lifestyle, and economic status. A metabolomics investigation was conducted on 8,299 unrelated subjects within the CLSA cohort, measuring their plasma metabolites and conducting genome-wide association analyses. This study represents the most recent and comprehensive genome-wide association study of metabolites,

encompassing data on 1,091 plasma metabolites and 309 plasma metabolite ratios. For further details on the GWAS, please refer to the original publication (Chen et al., 2023). The instrumental variables for T2DM were acquired through the IEU Open GWAS database (ID: ebi-a-GCST90018926). The dataset utilized in this study originated from Sakaue's meta-analysis, incorporating data from the United Kingdom Biobank and FinnGen (Sakaue et al., 2021). GWAS data for T2DM were extracted from this meta-analysis, which included a total of 24,167,560 SNPs identified across 4,90089 individuals of European ancestry. To the best of our knowledge, there is no sample overlap between exposures and outcomes in the current study, rendering the data relatively reliable.

2.4 Screening of instrumental variables

To identify qualified single-nucleotide polymorphisms (SNPs), we established a screening threshold. In this study, we adopted a relaxed association threshold of $p < 1 \times 10^{-5}$ (Xiao et al., 2022) to identify SNPs strongly associated with 1,400 exposures. Additionally, to mitigate the influence of linkage disequilibrium (LD), we implemented criteria of $r^2 < 0.001$ and a kilobase pair (kb) value of 10,000. To diminish the impact of weak instrumental variables (IVs) on the study's outcomes, we assessed completed SNPs using the F-statistic (Palmer et al., 2012). A threshold of F-statistic >10 indicated a low likelihood of the IV being weak (von Zglinicki, 2002). Furthermore, to mitigate the influence of additional potential confounders on the study outcome variables, we employed

the Phenotype Scanner V2.0 database (Kamat et al., 2019) to exclude SNPs correlated with factors such as smoking, alcohol consumption, physical activity, and education level.

2.5 Statistical analysis

We utilized the TwoSampleMR software package of R4.2.3 to analyze 1,091 blood metabolites and 309 metabolite ratios in relation to T2DM through MR. This study employed five commonly used analytical methods, namely, Inverse variance weighted (IVW), Weighted median, MR Egger, Weighted mode, and Simple mode, to determine their causal associations with T2DM.The IVW method, the primary approach in MR analysis, assumes the validity of each SNP, enabling a reliable assessment of exposure's causal effect on the outcome (Hartwig et al., 2017). MR Egger regression analysis accommodates potential pleiotropy or a substantial number of invalid instrumental variables for causal inference (Bowden et al., 2015). The Weighted median approach is utilized when at least 50% of valid instrumental variables are presumed to be present (Slob and Burgess, 2020). Weighted mode and Simple mode are alternate methods that relax assumptions, albeit with lower test efficacy compared to the previous three methods (Hartwig et al., 2017). These methods serve as supplementary tools for MR analysis. Additionally, to address multiple comparisons, False Discovery Rates (FDR) proposed by Benjamini and Hochberg were applied to correct p-values in the IVW results (Curtin and Schulz, 1998). A significance threshold of $p < 0.05$ after FDR correction was utilized.

2.6 Sensitivity analysis

To ensure result quality, a comprehensive sensitivity analysis was conducted using TwoSampleMR and MR-PRESSO software packages within R 4.2.3. Cochran's Q test assessed SNP heterogeneity (Greco et al., 2015). The MR-Egger intercept test and MR-PRESSO Global examined pleiotropy. Horizontal pleiotropy suggests a non-causal association pathway between IVs and outcomes, potentially causing false positives (Hemani et al., 2018). MR-PRESSO identified and excluded significant outliers, if present, followed by a reiteration of the MR analysis. Additionally, the Steiger test was applied to mitigate bias arising from reverse causality (Hemani et al., 2017). Individual SNP effect analyses and leave-one-out sensitivity analyses were performed to identify SNPs susceptible to significant heterogeneity.

2.7 Replication analysis

To further validate the associations identified in the primary analysis, we conducted replication analyses of metabolites and metabolite ratios initially found to exhibit significant causal associations (FDR <0.05) using independent GWAS data for blood metabolites. For this purpose, we utilized the IEU Open GWAS database (ID: ebi-a-GCST006867), which provided the largest sample size GWAS data available for T2DM, comprising 659,316 individuals of European ancestry. This dataset, derived

from a meta-analysis conducted by Xue et al. (2018), incorporated raw data primarily sourced from datasets such as the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) and Genetic Epidemiology Research on Adult Health and Aging (GERA), totaling 62,892 T2DM cases and 596,424 controls of European ancestry, and encompassing over five million genetic variants. For additional details regarding the GWAS, please refer to the original publication (Xue et al., 2018). The Mendelian randomization (MR) analysis methods employed in the replication analyses were consistent with those utilized in the main analyses and comprised the IVW, weighted median, MR Egger, weighted mode, and simple mode approaches, with the IVW results serving as the primary outcomes. Furthermore, a comprehensive sensitivity analysis was conducted to assess the robustness of the findings. Replication analyses were conducted to validate the reliability of our results.

2.8 Metabolic pathway analysis

Metabolic pathway analysis was conducted via MetaboAnalyst 5.0 for the metabolites exhibiting significant causal associations with T2DM identified through the IVW method in this study (https://www.metaboanalyst.ca/) (Chong et al., 2018). MetaboAnalyst 5. 0 serves as a comprehensive web-based data analysis tool designed to aid users in metabolomics data analysis and visualization. Utilizing this platform enables the identification of potential metabolite pathways associated with the underlying biological mechanisms of T2DM.

3 Results

3.1 Instrumental variables

After a meticulous series of screening steps, a final set of 1,091 blood metabolites and 309 metabolite ratios were obtained, featuring a range of 12–93 IVs. Among the 1,091 blood metabolites, IVs varied from 12 to 93, with X-15523 yielding the highest number of IVs and X-12462 the lowest. Similarly, the 309 metabolite ratios exhibited IVs ranging from 13 to 39, with the glutamine to alanine ratio presenting the highest number and the adenosine 5′-diphosphate to uridine ratio the fewest IVs. The F-statistic of the SNPs analyzed in this study ranged from 19.50 to 5,308.35, indicating a minimal likelihood of weak instrumental variables. These findings support the validity of all IVs for conducting Mendelian randomization analyses involving the 1,091 blood metabolites and 309 metabolite ratios (Supplementary Table S1).

3.2 Causal association of 1,091 blood metabolites and 309 metabolite ratios on T2DM

Causal associations of 1,091 blood metabolites and 309 metabolite ratios for T2DM were determined by 5 MR analysis methods (IVW, Weighted median, MR Egger, Weighted mode and Simple mode). We identified a total of 185 blood metabolites and 70 metabolite ratios ($p < 0.05$ for the presence of
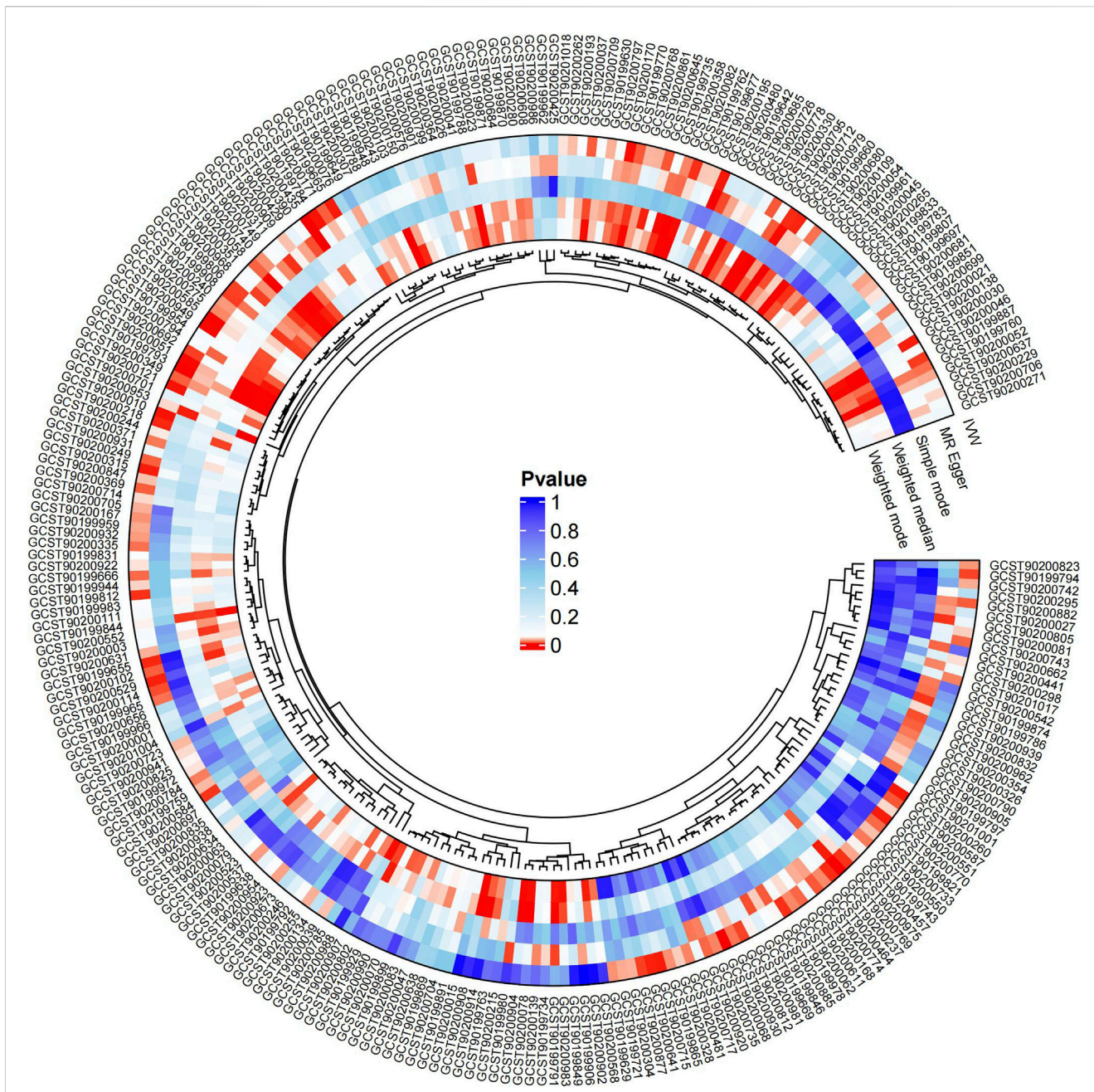
FIGURE 2
Circos plot of MR analysis results for 185 blood metabolites and 70 metabolite ratios.

at least 1 MR analysis method) with a significant causal association with T2DM (Supplementary Table S2). Visualize this result with a circos plot (Figure 2). The figure comprises 1,275 color blocks, each representing an individual MR analysis result for one of the exposures. Different colors denote various p-values, with shades closer to dark blue indicating larger p-values and those closer to red indicating smaller p-values. The IVW analysis identified 88 blood metabolites and 37 metabolite ratios as having a significant causal relationship with T2DM ($p < 0.05$) (Supplementary Table S3). It contains 74 known metabolites. After correction based on the FDR method, a total of three known blood metabolites, one unknown blood metabolite, and one metabolite ratios were identified as having

a significant causal relationship with T2DM ($p < 0.05$) (Supplementary Table S4). They are respectively: 1-linoleoyl-GPE (18:2) (IVW: OR:0.930, 95%CI: 0.899–0.962, $p = 2.16 \times 10^{-5}$, FDR = 0.008), 1,2-dilinoleoyl-GPE (18:2/18:2) (IVW: OR:0.942, 95%CI: 0.917–0.968, $p = 1.64 \times 10^{-5}$, FDR = 0.008), Mannose (IVW: OR: 1.133, 95%CI: 1.072–1.197, $p = 1.02 \times 10^{-5}$, FDR = 0.014), X-21829 (IVW: OR:1.036, 95%CI:1.036–1.122, $p = 9.44 \times 10^{-5}$, FDR = 0.026), Phosphate to mannose ratio (IVW: OR:0.870, 95%CI: 0.818–0.926, $p = 1.29 \times 10^{-5}$, FDR = 0.008). This result is visualized through a forest plot (Figure 3). The bolded p-values in the figure indicate statistical significance, while the five colored nodes represent five distinct MR analysis methods. Among them 1-linoleoyl-GPE (18:2),

exposure	nsnp	method	pval		OR(95% CI)
1-linoleoyl-GPE (18:2)	33	MR Egger	**0.034**		0.930 (0.872 to 0.992)
	33	Weighted median	**<0.001**		0.910 (0.877 to 0.945)
	33	Inverse variance weighted	**<0.001**		0.930 (0.899 to 0.962)
	33	Simple mode	**0.018**		0.910 (0.845 to 0.980)
	33	Weighted mode	**<0.001**		0.910 (0.872 to 0.950)
1,2-dilinoleoyl-GPE (18:2/18:2)	28	MR Egger	0.122		0.954 (0.900 to 1.011)
	28	Weighted median	0.266		0.977 (0.936 to 1.018)
	28	Inverse variance weighted	**<0.001**		0.942 (0.917 to 0.968)
	28	Simple mode	0.755		0.987 (0.908 to 1.072)
	28	Weighted mode	0.614		0.981 (0.910 to 1.057)
Mannose	26	MR Egger	**<0.001**		1.288 (1.177 to 1.410)
	26	Weighted median	**0.014**		1.085 (1.017 to 1.158)
	26	Inverse variance weighted	**<0.001**		1.133 (1.072 to 1.197)
	26	Simple mode	0.170		1.075 (0.972 to 1.188)
	26	Weighted mode	0.135		1.070 (0.982 to 1.166)
X-21829	30	MR Egger	**0.041**		1.101 (1.008 to 1.202)
	30	Weighted median	**0.009**		1.070 (1.017 to 1.126)
	30	Inverse variance weighted	**<0.001**		1.074 (1.036 to 1.112)
	30	Simple mode	0.090		1.091 (0.990 to 1.203)
	30	Weighted mode	0.074		1.073 (0.996 to 1.157)
Phosphate to mannose	26	MR Egger	**0.001**		0.773 (0.672 to 0.889)
	26	Weighted median	**0.010**		0.913 (0.852 to 0.978)
	26	Inverse variance weighted	**<0.001**		0.870 (0.818 to 0.926)
	26	Simple mode	0.320		0.941 (0.836 to 1.058)
	26	Weighted mode	0.572		0.945 (0.780 to 1.146)

FIGURE 3
Forest plot of FDR-corrected significant results.

1,2-dilinoleoyl-GPE (18:2/18:2) and Phosphate to mannose ratio were protective factors for T2DM and X-21829 and Mannose were risk factors for T2DM. This result is further visualized by means of a scatterplot (Figure 4). The graph displays five distinct colored line segments, each corresponding to a different method of MR analysis. The slope of the lines indicates the direction of the causal association, whether positive or negative.

3.3 Sensitive analysis

We conducted a sensitivity analysis to evaluate the stability of the five exposures that exhibited IVW results consistent with FDR<0.05 correction. The Cochran Q test indicated no significant heterogeneity between the SNPs of the IVW method and the MR-Egger method ($p > 0.05$) (Supplementary Table S5). Both the MR-Egger intercept test and the MR-PRESSO Global test results suggested the absence of horizontal pleiotropy in our study ($p > 0.05$) (Supplementary Table S6). Additionally, the MR-PRESSO outlier test did not detect significant outlier SNPs. Our leave-one-out sensitivity analyses and the MR analyses of individual SNPs (Supplementary Figures S1–S10) demonstrated the robustness of

our MR analyses, with no instances of individual SNPs significantly influencing the outcomes. Moreover, the Steiger test for directionality resulted in TRUE ($p < 0.05$) in our study. Notably, the funnel plot illustrated a generally symmetrical distribution of IVW results in this study without notable bias (Supplementary Figures S11–S15). These findings collectively contribute to the increased reliability of the results obtained in this study.

3.4 Replication analysis

To bolster the credibility of our study's findings, we conducted a replication analysis on the five exposures meeting the FDR<0.05 correction criteria using an alternative GWAS dataset for T2DM (Xue et al., 2018). Encouragingly, our observations within this GWAS outcome data aligned with similar trends for these five exposures. However, the unknown metabolite X-21829 did not exhibit a significant difference concerning T2DM. Among these, three metabolites: 1-linoleoyl-GPE (18:2) (IVW: OR:0.945, 95%CI: 0.904–0.989, $p = 0.015$), 1,2-dilinoleoyl-GPE (18:2/18:2) (IVW: OR: 0.925, 95%CI: 0.870–0.984, $p = 0.014$), Mannose (IVW: OR:1.143, 95%CI: 1.053–1.241, $p = 0.001$), and Phosphate to mannose ratio
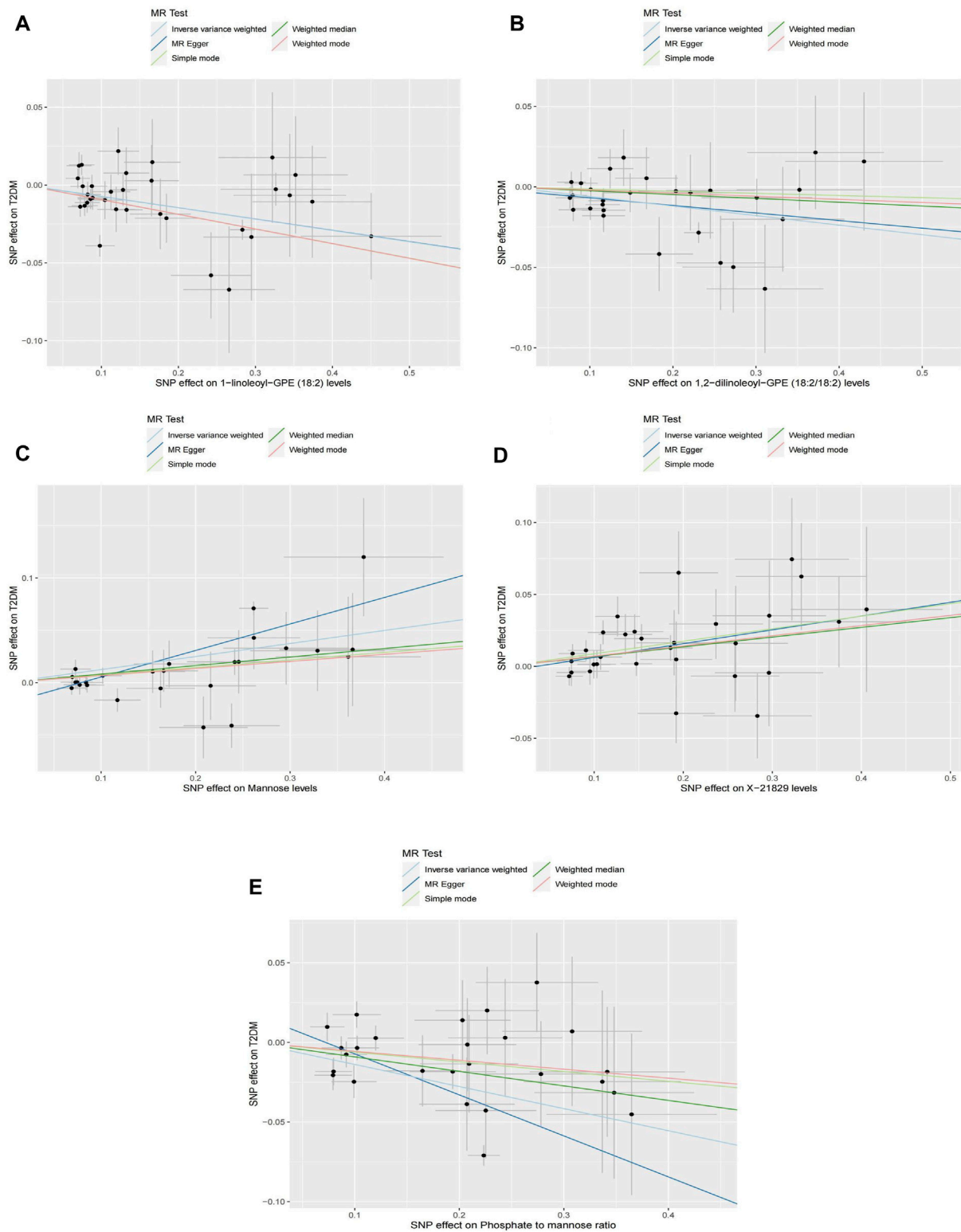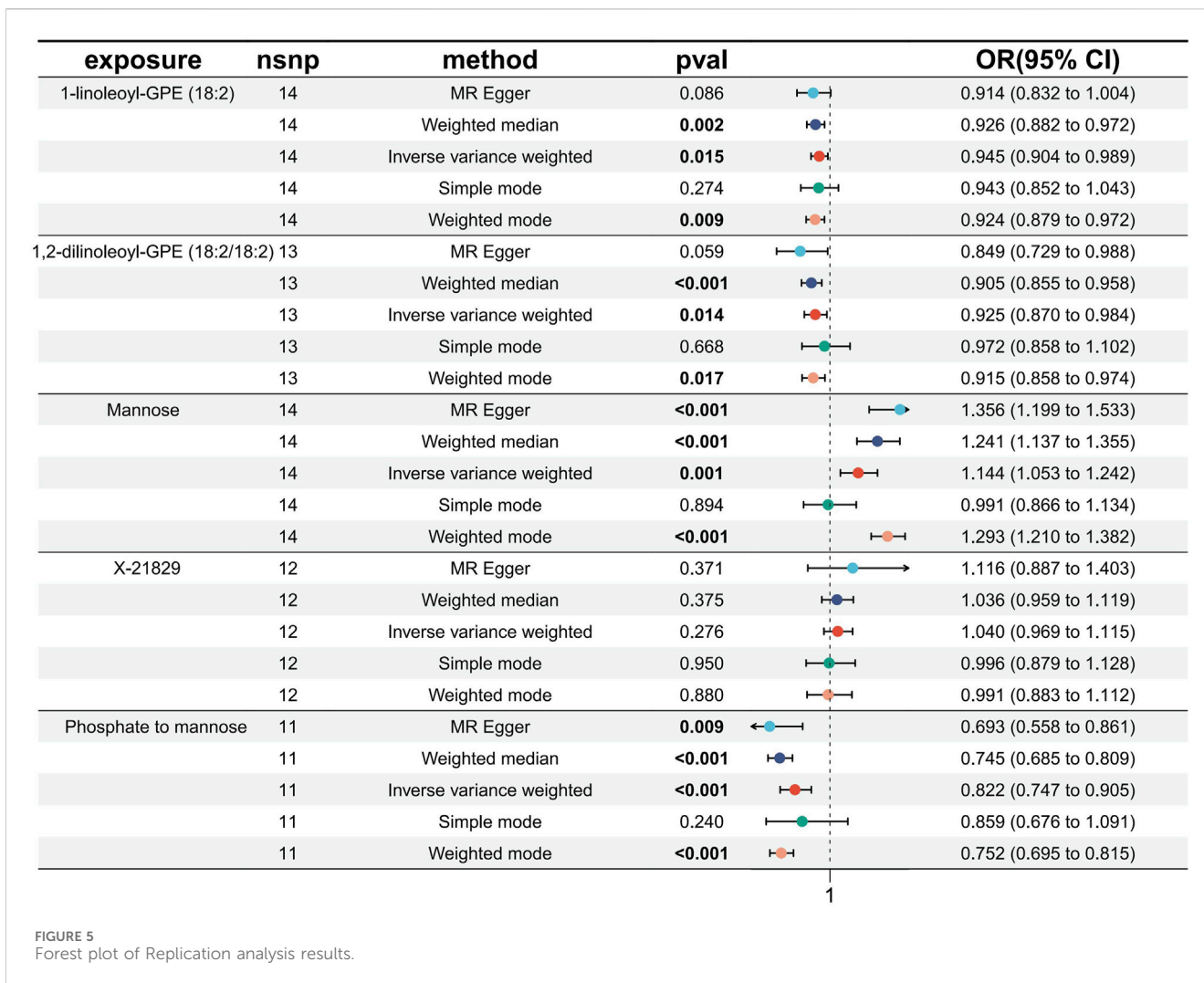
FIGURE 4
Scatterplot of causal associations between core metabolites and Phosphate to mannose on T2DM Note: **(A)** 1-linoleoyl-GPE (18:2) on T2DM. **(B)** 1,2-dilinoleoyl-GPE (18:2/18:2) on T2DM. **(C)** Mannose on T2DM. **(D)** X-21829 on T2DM. **(E)** Phosphate to mannose on T2DM.

(IVW: OR:0.822, 95%CI: 0.747–0.905, $p = 6.64 \times 10^{-5}$) maintained their significant causal associations with T2DM (Figure 5). The bolded p-values in the figure indicate statistical significance, while

the five colored nodes represent five distinct MR analysis methods. The confidence in these four metabolites and their ratios remains high.

exposure	nsnp	method	pval		OR(95% CI)
1-linoleoyl-GPE (18:2)	14	MR Egger	0.086		0.914 (0.832 to 1.004)
	14	Weighted median	**0.002**		0.926 (0.882 to 0.972)
	14	Inverse variance weighted	**0.015**		0.945 (0.904 to 0.989)
	14	Simple mode	0.274		0.943 (0.852 to 1.043)
	14	Weighted mode	**0.009**		0.924 (0.879 to 0.972)
1,2-dilinoleoyl-GPE (18:2/18:2)	13	MR Egger	0.059		0.849 (0.729 to 0.988)
	13	Weighted median	**<0.001**		0.905 (0.855 to 0.958)
	13	Inverse variance weighted	**0.014**		0.925 (0.870 to 0.984)
	13	Simple mode	0.668		0.972 (0.858 to 1.102)
	13	Weighted mode	**0.017**		0.915 (0.858 to 0.974)
Mannose	14	MR Egger	**<0.001**		1.356 (1.199 to 1.533)
	14	Weighted median	**<0.001**		1.241 (1.137 to 1.355)
	14	Inverse variance weighted	**0.001**		1.144 (1.053 to 1.242)
	14	Simple mode	0.894		0.991 (0.866 to 1.134)
	14	Weighted mode	**<0.001**		1.293 (1.210 to 1.382)
X-21829	12	MR Egger	0.371		1.116 (0.887 to 1.403)
	12	Weighted median	0.375		1.036 (0.959 to 1.119)
	12	Inverse variance weighted	0.276		1.040 (0.969 to 1.115)
	12	Simple mode	0.950		0.996 (0.879 to 1.128)
	12	Weighted mode	0.880		0.991 (0.883 to 1.112)
Phosphate to mannose	11	MR Egger	**0.009**		0.693 (0.558 to 0.861)
	11	Weighted median	**<0.001**		0.745 (0.685 to 0.809)
	11	Inverse variance weighted	**<0.001**		0.822 (0.747 to 0.905)
	11	Simple mode	0.240		0.859 (0.676 to 1.091)
	11	Weighted mode	**<0.001**		0.752 (0.695 to 0.815)

1

FIGURE 5
Forest plot of Replication analysis results.

3.5 Metabolic pathway analysis

Pathway analysis of 74 known metabolites with statistically significant ($p < 0.05$) IVW results unveiled six metabolic pathways with considerable significance ($p < 0.05$): Valine, leucine, and isoleucine biosynthesis ($p = 0.004$), Phenylalanine metabolism ($p = 0.007$), Glycerophospholipid metabolism ($p = 0.010$), Alpha-Linolenic acid metabolism ($p = 0.011$), Sphingolipid metabolism ($p = 0.029$), and Alanine, aspartate, and glutamate metabolism ($p = 0.049$) (Supplementary Table S7) (Figure 6). The nodes in the graph depict metabolic pathways, wherein nodes that are higher and darker indicate smaller p-values associated with the respective metabolic pathway.

4 Discussion

The substantial global burden of high morbidity and mortality rates associated with T2DM has underscored the urgent need for early screening and prevention measures (Collins et al., 2011; Mulnier et al., 2006; Ogurtsova et al., 2022). Achieving high-quality medical care necessitates tailored and precise treatments,

crucial for foreseeing individual health indicators and preventing the onset of T2DM(Hampel et al., 2021). Biomarkers serve as pivotal tools in disease management and prevention. Personalized and multidimensional biomarkers significantly contribute to predicting, diagnosing, and prognosing T2DM, offering invaluable insights for drug development, clinical diagnosis, and individualized treatment strategies. The advancement of histological technology facilitates understanding the molecular mechanisms underlying T2DM and evaluating biomarkers, thereby significantly advancing the realm of precision medicine for T2DM.

This study represents an extensive investigation into T2DM, integrating genomics and metabolomics. We utilized GWAS data from two large-scale T2DM cohorts to examine the causal relationship between 1,091 blood metabolites and 309 metabolite ratios with T2DM. Initially, employing IVW, Weighted Median, MR Egger, Weighted Mode, and Simple Mode methods, we identified 185 metabolites and 70 metabolite ratios with at least one method showing significant results ($p < 0.05$). Specifically, the IVW analysis revealed 88 blood metabolites and 37 metabolite ratios displaying a significant causal association with T2DM ($p < 0.05$), encompassing 74 known metabolites. From these 74 metabolites, six metabolic pathways potentially involved in T2DM pathogenesis were

FIGURE 6
Metabolic pathway analysis bubble diagram.
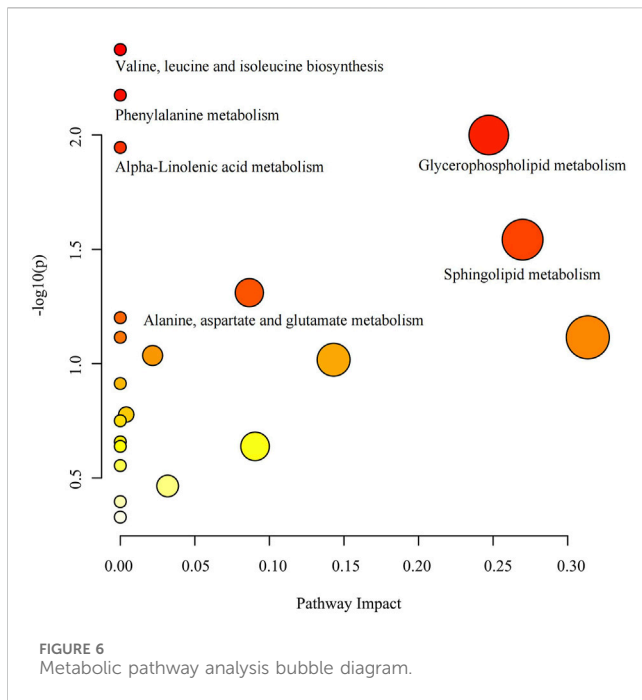
identified: Valine, leucine, and isoleucine biosynthesis ($p = 0.004$), Phenylalanine metabolism ($p = 0.007$), Glycerophospholipid metabolism ($p = 0.010$), Alpha-Linolenic acid metabolism ($p = 0.011$), Sphingolipid metabolism ($p = 0.029$), and Alanine, aspartate, and glutamate metabolism ($p = 0.049$). Additionally, through FDR multiple test correction, we identified three known metabolites and one metabolite ratio: 1-linoleoyl-GPE (18:2), 1,2-dilinoleoyl-GPE (18:2/18:2), Mannose, and the Phosphate to Mannose ratio. Subsequently, we validated these findings using another GWAS dataset for T2DM with a substantially larger sample size (Xue et al., 2018), and fortunately, these three known metabolites and one metabolite ratio remained significantly associated in this independent dataset.

The exposure data used in this study originated from Chen's GWAS investigation (Chen et al., 2023), which stands as the most recent and comprehensive study incorporating metabolites. Consequently, the metabolites under examination are relatively novel, with many of their connections to T2DM yet to be explored. Among our findings, the most significant causal link identified was between 1-linoleoyl-GPE (18:2) and T2DM. Elevated levels of 1-linoleoyl-GPE (18:2) were notably associated with a reduced risk of developing T2DM. However, it is important to note that 1-linoleoyl-GPE (18:2) lacks prior studies exploring its correlation with T2DM. Nevertheless, research has highlighted the significance of linoleoylethanolamide in ameliorating weight gain, dyslipidemia, and inflammation induced by a high-fat diet (Tovar et al., 2023). Considering the established association between obesity, fat accumulation, insulin resistance, and T2DM, it is plausible that 1-linoleoyl-GPE (18:2) may act as a protective factor. Another significant protective metabolite identified in our study is 1,2-dilinoleoyl-GPE (18:2/18:2), yet there is a lack of studies exploring its relevance to T2DM. Jansen's team (Jansen et al., 2023) observed that 1,2-Dilinoleoyl-sn-glycero-3-phosphocholine enhances adipocyte catabolism and apoptosis through a TNF-α-dependent

pathway, thereby alleviating insulin resistance *via* PPARα-mediated inhibition of myocyte inflammation. Furthermore, our study has identified Mannose as a risk factor for the development of T2DM. Mannose, a crucial hexose for glycoprotein synthesis, consistently demonstrates a significant association with elevated blood glucose levels and T2DM development in prospective studies (Carter et al., 2016; Floegel et al., 2013; Menni et al., 2013; Sone et al., 2003). This relationship might elucidate the Phosphate to Mannose ratio, where Mannose, as the denominator, potentially serves as a protective factor against T2DM. Among the results obtained from metabolic pathway analysis, the Valine, leucine, and isoleucine biosynthesis pathways exhibited the highest significance ($p = 0.004$). Valine, leucine, and isoleucine within this metabolic pathway act as vital nutritional signals influencing protein synthesis, glucose regulation, and anti-obesity mechanisms (Nie et al., 2018). Specifically, isoleucine stimulates glucose uptake in skeletal muscle, thereby preventing spikes in plasma glucose concentrations, while also exhibiting a preventive effect against visceral obesity and hyperinsulinemia (Doi et al., 2005; Nishimura et al., 2010).

Our study exhibits several strengths. Firstly, we leveraged the most advanced and comprehensive GWAS data encompassing 1,091 blood metabolites and 309 metabolite ratios. Consequently, our study stands as the most cutting-edge exploration of causal associations between metabolites and T2DM to date. Moreover, we implemented a rigorous MR study methodology to mitigate confounding factors commonly present in observational studies. To ensure the robustness of our findings, we conducted a series of sensitivity analyses, enhancing the reliability of our results. Additionally, we validated our findings using an alternate, larger sample set from a separate GWAS dataset for T2DM, obtaining consistently aligned results with our original study. Lastly, our utilization of metabolic pathway analyses offers insights into the metabolic mechanisms associated with the onset of T2DM, providing valuable reference points for further investigation.

Nevertheless, this study bears certain limitations. Initially, our inclusion covered a relatively restricted subset of the 1,400 exposed SNPs, thereby necessitating a more permissive threshold during the screening of instrumental variables for MR analysis, akin to other studies of a similar nature. Additionally, a limited number of metabolites were inevitably omitted from pathway analysis due to lacking nomenclature or annotations in the metabolic pathway database. To address these limitations, future investigations should prioritize further experimentation on less explored metabolites, offering a more comprehensive understanding of their association with T2DM. Moreover, expanding the sample size of the original dataset and conducting high-quality randomized controlled trials alongside fundamental studies will be pivotal for validating our findings in subsequent research.

5 Conclusion

To summarize, this MR analysis unveiled 88 blood metabolites and 37 metabolite ratios exhibiting significant causal links to T2DM. Furthermore, through FDR validation, three established metabolites and one metabolite ratio emerged as displaying the most robust causal association with T2DM. Additionally, our investigation identified six metabolic pathways potentially linked to T2DM

development. These discerned serum metabolites establish a foundation for early screening, preventive strategies, and treatment protocols for T2DM, while also guiding the blueprint for future clinical studies. The amalgamation of genomics and metabolomics in MR analysis serves as a pivotal pathway for delving into the etiology and pathogenesis of T2DM.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

JL: Conceptualization, Data curation, Formal Analysis, Resources, Writing–original draft. WW: Conceptualization, Data curation, Formal Analysis, Methodology, Software, Writing–review and editing. FL: Conceptualization, Data curation, Software, Writing–review and editing. LQ: Investigation, Methodology, Supervision, Writing–original draft. YR: Methodology, Software, Writing–review and editing. ML: Validation, Writing–review and editing. WL: Validation, Writing–review and editing. FG: Funding acquisition, Writing–review and editing. JZ: Funding acquisition, Writing–review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1356696/full#supplementary-material

References

Bauermeister, A., Mannochio-Russo, H., Costa-Lotufo, L. V., Jarmusch, A. K., and Dorrestein, P. C. (2022). Mass spectrometry-based metabolomics in microbiome investigations. *Nat. Rev. Microbiol.* 20(3), 143-160. doi:10.1038/s41579-021-00621-9

Bloomgarden, Z. (2018). Diabetes and branched-chain amino acids: what is the link?. *J. Diabetes* 10 (5), 350–352. doi:10.1111/1753-0407.12645

Bowden, J., Davey, S. G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44 (2), 512–525. doi:10.1093/ije/dyv080

Carter, T. C., Rein, D., Padberg, I., Peter, E., Rennefahrt, U., David, D. E., et al. (2016). Validation of a metabolite panel for early diagnosis of type 2 diabetes. *Metabolism* 65 (9), 1399–1408. doi:10.1016/j.metabol.2016.06.007

Chen, Y., Lu, T., Pettersson-Kymmer, U., Stewart, I. D., Butler-Laporte, G., Nakanishi, T., et al. (2023). Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. *Nat. Genet.* 55 (1), 44–53. doi:10.1038/s41588-022-01270-1

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 46 (W1), W486–W494. doi:10.1093/nar/gky310

Collins, G. S., Mallett, S., Omar, O., and Yu, L. M. (2011). Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med.* 9, 103. doi:10.1186/1741-7015-9-103

Curtin, F., and Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biol. Psychiatry* 44 (8), 775–777. doi:10.1016/s0006-3223(98)00043-2

Doi, M., Yamaoka, I., Nakayama, M., Mochizuki, S., Sugahara, K., and Yoshizawa, F. (2005). Isoleucine, a blood glucose-lowering amino acid, increases glucose uptake in rat skeletal muscle in the absence of increases in AMP-activated protein kinase activity. *J. Nutr.* 135 (9), 2103–2108. doi:10.1093/jn/135.9.2103

Felig, P., Marliss, E., and Cahill, G. J. (1969). Plasma amino acid levels and insulin secretion in obesity. *N. Engl. J. Med.* 281 (15), 811–816. doi:10.1056/NEJM196910092811503

Floegel, A., Stefan, N., Yu, Z., Muhlenbruch, K., Drogan, D., Joost, H. G., et al. (2013). Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 62 (2), 639–648. doi:10.2337/db12-0495

Greco, M. F., Minelli, C., Sheehan, N. A., and Thompson, J. R. (2015). Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* 34 (21), 2926–2940. doi:10.1002/sim.6522

Hampel, H., Nistico, R., Seyfried, N. T., Levey, A. I., Modeste, E., Lemercier, P., et al. (2021). Omics sciences for systems biology in Alzheimer's disease: State-of-the-art of the evidence. *Ageing Res. Rev.* 69, 101346. doi:10.1016/j.arr.2021.101346

Hartwig, F. P., Davey, S. G., and Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* 46 (6), 1985–1998. doi:10.1093/ije/dyx102

Hemani, G., Tilling, K., and Davey, S. G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 13 (11), e1007081. doi:10.1371/journal.pgen.1007081

Hemani, G., Bowden, J., and Davey, S. G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* 27 (R2), R195–R208. doi:10.1093/hmg/ddy163

Holland, W. L., Bikman, B. T., Wang, L. P., Yuguang, G., Sargent, K. M., Bulchand, S., et al. (2011). Lipid-induced insulin resistance mediated by the proinflammatory receptor TLR4 requires saturated fatty acid-induced ceramide biosynthesis in mice. *J. Clin. Invest.* 121 (5), 1858–1870. doi:10.1172/JCI43378

Imamura, F., Micha, R., Wu, J. H., de Oliveira, O. M., Otite, F. O., Abioye, A. I., et al. (2016). Effects of saturated fat, polyunsaturated fat, monounsaturated fat, and carbohydrate on glucose-insulin homeostasis: a systematic review and meta-analysis of randomised controlled feeding trials. *PLoS Med.* 13 (7), e1002087. doi:10.1371/journal.pmed.1002087

Jansen, R., Milaneschi, Y., Schranner, D., Kastenmuller, G., Arnold, M., Han, X., et al. (2023). The Metabolome-Wide signature of major depressive disorder. *Res. Sq.* doi:10.21203/rs.3.rs-3127544/v1

Johnson, C. H., Ivanisevic, J., and Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* 17 (7), 451–459. doi:10.1038/nrm.2016.25

Kamat, M. A., Blackshaw, J. A., Young, R., Surendran, P., Burgess, S., Danesh, J., et al. (2019). PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 35 (22), 4851–4853. doi:10.1093/bioinformatics/btz469

Keller-Pinter, A., Korcsmaros, T., and Vellai, T. (2023). Managing type 2 diabetes: Targeting a microbial enzyme as a novel treatment option. *Signal Transduct. Target Ther.* 8 (1), 444. doi:10.1038/s41392-023-01694-z

Long, J., Yang, Z., Wang, L., Han, Y., Peng, C., Yan, C., et al. (2020). Metabolite biomarkers of type 2 diabetes mellitus and pre-diabetes: a systematic review and meta-analysis. *BMC Endocr. Disord.* 20 (1), 174. doi:10.1186/s12902-020-00653-x

Magliano, D. J., Sacre, J. W., Harding, J. L., Gregg, E. W., Zimmet, P. Z., and Shaw, J. E. (2020). Young-onset type 2 diabetes mellitus - implications for morbidity and mortality. *Nat. Rev. Endocrinol.* 16 (6), 321–331. doi:10.1038/s41574-020-0334-z

Menni, C., Fauman, E., Erte, I., Perry, J. R., Kastenmuller, G., Shin, S. Y., et al. (2013). Biomarkers for type 2 diabetes and impaired fasting glucose using a nontargeted metabolomics approach. *Diabetes* 62 (12), 4270–4276. doi:10.2337/db13-0570

Mora-Ortiz, M., Alcala-Diaz, J. F., Rangel-Zuniga, O. A., Arenas-de, L. A., Abollo-Jimenez, F., Luque-Cordoba, D., et al. (2022). Metabolomics analysis of type 2 diabetes remission identifies 12 metabolites with predictive capacity: a CORDIOPREV clinical trial study. *BMC Med.* 20 (1), 373. doi:10.1186/s12916-022-02566-z

Mulnier, H. E., Seaman, H. E., Raleigh, V. S., Soedamah-Muthu, S. S., Colhoun, H. M., and Lawrenson, R. A. (2006). Mortality in people with type 2 diabetes in the UK. *Diabet. Med.* 23 (5), 516–521. doi:10.1111/j.1464-5491.2006.01838.x

Newgard, C. B., An, J., Bain, J. R., Muehlbauer, M. J., Stevens, R. D., Lien, L. F., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* 9 (4), 311–326. doi:10.1016/j.cmet.2009.02.002

Nie, C., He, T., Zhang, W., Zhang, G., and Ma, X. (2018). Branched chain amino acids: beyond nutrition metabolism. *Int. J. Mol. Sci.* 19 (4), 954. doi:10.3390/ijms19040954

Nishimura, J., Masaki, T., Arakawa, M., Seike, M., and Yoshimatsu, H. (2010). Isoleucine prevents the accumulation of tissue triglycerides and upregulates the expression of PPARalpha and uncoupling protein in diet-induced obese mice. *J. Nutr.* 140 (3), 496–500. doi:10.3945/jn.109.108977

Ogurtsova, K., Guariguata, L., Barengo, N. C., Ruiz, P. L., Sacre, J. W., Karuranga, S., et al. (2022). IDF diabetes Atlas: global estimates of undiagnosed diabetes in adults for 2021. *Diabetes Res. Clin. Pract.* 183, 109118. doi:10.1016/j.diabres.2021.109118

Palmer, T. M., Lawlor, D. A., Harbord, R. M., Sheehan, N. A., Tobias, J. H., Timpson, N. J., et al. (2012). Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat. Methods Med. Res.* 21 (3), 223–242. doi:10.1177/0962280210394459

Palmer, N. D., Stevens, R. D., Antinozzi, P. A., Anderson, A., Bergman, R. N., Wagenknecht, L. E., et al. (2015). Metabolomic profile associated with insulin resistance and conversion to diabetes in the Insulin Resistance Atherosclerosis Study. *J. Clin. Endocrinol. Metab.* 100 (3), E463–E468. doi:10.1210/jc.2014-2357

Parcha, V., Heindl, B., Kalra, R., Li, P., Gower, B., Arora, G., et al. (2022). Insulin resistance and cardiometabolic risk profile among nondiabetic american young adults: insights from NHANES. *J. Clin. Endocrinol. Metab.* 107 (1), e25–e37. doi:10.1210/clinem/dgab645

Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., et al. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* 53 (10), 1415–1424. doi:10.1038/s41588-021-00931-x

Siddik, M., and Shin, A. C. (2019). Recent progress on branched-chain amino acids in obesity, diabetes, and beyond. *Endocrinol. Metab. Seoul.* 34 (3), 234–246. doi:10.3803/EnM.2019.34.3.234

Skrivankova, V. W., Richmond, R. C., Woolf, B., Yarmolinsky, J., Davies, N. M., Swanson, S. A., et al. (2021). Strengthening the reporting of observational studies in Epidemiology using mendelian randomization: the STROBE-MR statement. *JAMA* 326 (16), 1614–1621. doi:10.1001/jama.2021.18236

Slob, E., and Burgess, S. (2020). A comparison of robust Mendelian randomization methods using summary data. *Genet. Epidemiol.* 44 (4), 313–329. doi:10.1002/gepi.22295

Sone, H., Shimano, H., Ebinuma, H., Takahashi, A., Yano, Y., Iida, K. T., et al. (2003). Physiological changes in circulating mannose levels in normal, glucose-intolerant, and diabetic subjects. *Clin. Trial Metabolism.* 52 (8), 1019–1027. doi:10.1016/s0026-0495(03)00153-7

Tovar, R., de Ceglia, M., Ubaldi, M., Rodriguez-Pozo, M., Soverchia, L., Cifani, C., et al. (2023). Administration of linoleoylethanolamide reduced weight gain, dyslipidemia, and inflammation associated with High-Fat-Diet-Induced obesity. *Nutrients* 15 (20), 4448. doi:10.3390/nu15204448

von Zglinicki, T. (2002). Oxidative stress shortens telomeres. *Trends biochem. Sci.* 27 (7), 339–344. doi:10.1016/s0968-0004(02)02110-2

Wang, Q., Holmes, M. V., Davey, S. G., and Ala-Korpela, M. (2017). Genetic support for a causal role of insulin resistance on circulating Branched-Chain amino acids and inflammation. *Diabetes Care* 40 (12), 1779–1786. doi:10.2337/dc17-1642

Xiao, G., He, Q., Liu, L., Zhang, T., Zhou, M., Li, X., et al. (2022). Causality of genetically determined metabolites on anxiety disorders: a two-sample Mendelian randomization study. *J. Transl. Med.* 20 (1), 475. doi:10.1186/s12967-022-03691-2

Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., et al. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat. Commun.* 9 (1), 2941. doi:10.1038/s41467-018-04951-w

Yang, Q., Vijayakumar, A., and Kahn, B. B. (2018). Metabolites as regulators of insulin sensitivity and metabolism. *Nat. Rev. Mol. Cell Biol.* 19 (10), 654–672. doi:10.1038/s41580-018-0044-8