Diabetology & Metabolic Syndrome

RESEARCH

Open Access

Causal roles of lipids and mediating proteins in diabetic retinopathy: insights from metabolomic and proteomic Mendelian randomization

Jiawei Wang[1], Jing Su[1], Danyan Liu[1*†] and Jingxue Ma[1*†]

Abstract

Background This study explores the causal relationships between five major lipids, 249 circulating metabolites, and four diabetic retinopathy (DR) outcomes: overall DR, background DR, severe background DR, and proliferative DR (PDR). We aim to identify plasma proteins that mediate these causal effects, offering insights into potential therapeutic targets.

Methods We conducted metabolome-wide Mendelian randomization (MR) analyses to assess associations between major lipids, metabolites, and DR outcomes. Multivariable MR (MVMR) and proteome-wide mediated MR (two-step MR) analyses were performed to ensure robust evaluation and identify mediating plasma proteins.

Results Triglycerides were identified as a significant risk factor for DR, mediated by proteins like Dickkopf-3 (DKK3), ST6 N-acetylglucosamine transferase 6 (ST4S6), and Neogenin (NEO1). For background DR, HDL-C, specific VLDL particles, and LDL triglycerides were protective, mediated by proteins like chloride intracellular channel 5 (CLIC5), basal cell adhesion molecule (BCAM), and Ribophorin I (RPN1). Additionally, polyunsaturated fatty acids (PUFAs) and total choline were protective against PDR, mediated by Radical Fringe Gene (RFNG).

Conclusions This study identifies specific plasma proteins that mediate the effects of lipids and metabolites on DR, establishing a direct molecular link between these biomarkers and disease progression. These findings enhance our understanding of the pathophysiological mechanisms underlying DR and highlight potential targets for therapeutic intervention.

Keywords Diabetic retinopathy, Major lipids, Circulating metabolites, Proteome, causality

†Danyan Liu, Jingxue Ma co-corresponding author.

*Correspondence:
Danyan Liu
liudanyan@hebmu.edu.cn
Jingxue Ma
majingxue2003@163.com
[1]Department of Ophthalmology, The Second Hospital of Hebei Medical University, Shijiazhuang 050000, Hebei, China

Introduction

Diabetic retinopathy (DR) is the most common complication of diabetes mellitus (DM) and a leading cause of preventable blindness among adults, posing a significant public health challenge [1, 2]. By 2030, it is projected that 129.84 million adults globally will suffer from DR, with this number rising to 160.50 million by 2045 [3]. DR is characterized by progressive damage to the retinal microvasculature, with its development and progression are closely linked to metabolic dysregulation, including hyperglycemia and associated metabolic disturbances. Although elevated blood glucose levels has been recognized to contribute to the development of DR, differences in HbA1c levels explain only 6.6% of the variation in DR risk for the entire study cohort in a diabetes control and complications trial [4].

Dyslipidemia, defined by the presence of abnormal levels of lipids in the blood, is a common feature in individuals with diabetes and has been implicated in various diabetic complications, including DR [5]. Although epidemiological studies hint at a link between lipid irregularities and DR, establishing causality is complex, primarily due to the influence of confounding variables and the possibility of reverse causation [6]. Meta-analyses of DR research have identified associations between blood pressure, serum total cholesterol, and glycosylated hemoglobin levels with the prevalence of retinopathy. However, these factors collectively account for only a modest 9% of DR progression [7]. Metabolomics, a burgeoning field in the biomedical sciences, offers a novel perspective in the quest to understand complex diseases like DR. It is estimated that genetic variation accounts for approximately 50% of the observed phenotypic variance at the metabolite level, offering a unique opportunity for inferring causality from metabolite levels to disease risk [8]. Mendelian randomization (MR) represents a powerful approach in this context, utilizing genetic variants associated with a specific exposure to assess its causal impact on an outcome. By leveraging the principles of MR, researchers can navigate around the complexities of confounding factors and reverse causation, thereby enhancing the reliability of causality evidence in disease research. This method holds particular promise for elucidating the intricate interplay between metabolic pathways and DR pathogenesis.

In this study, we aimed to explore the causal associations between five major lipids and 249 circulating metabolites with various DR outcomes, including overall DR, background DR, severe background DR, and proliferative DR (PDR). By employing metabolome-wide MR analyses, we assessed the relationships between these factors and DR outcomes. Additionally, we utilized proteome-wide mediated MR analyses to identify plasma proteins that mediate these causal effects. Our findings provide valuable insights into the roles of specific lipids, metabolites, and proteins in DR, highlighting potential targets for therapeutic intervention and guiding future research directions.

Method

Data sources

The summary datasets used in this study are publicly available and can be accessed through the cited papers. All original GWAS studies included in this research were conducted with approval from their respective ethics committees, and informed consent was obtained from all participants involved in these studies.

This metabolome-wide MR study utilized publicly accessible summary datasets, as detailed in Table S1. In two-sample MR analyses, we obtained GWAS summary statistics for 254 circulating metabolites from the MRC-IEU OpenGWAS project (https://gwas.mrcieu.ac.

uk). These included data for five major non-fasted lipoprotein lipid traits measured using standard clinical chemistry assays in approximately 441,016 participants from the UK Biobank (UKBB) [9], as well as 249 metabolic biomarkers using high-throughput NMR spectroscopy in over 114,000 participants of European ancestry from the UKBB [10]. Summary-level data for outcomes, including diabetic retinopathy (DR), background DR, severe background DR, and proliferative diabetic retinopathy (PDR), were sourced from the FinnGen Release 9, a comprehensive European consortium (https://storage.googleapis.com/.

finngen-public-data-r9/summary_stats/) [11]. These datasets included 10,413 cases of DR compared to 308,633 control subjects, 4,011 background DR cases juxtaposed against 344,569 controls, 816 severe background DR cases compared to 344,569 controls, as well as 9,511 PDR cases against 362,581 controls. In the mediation analyses, we incorporated additional datasets beyond the previously described exposure and outcome data. Specifically, we utilized 4,489 GWAS summary statistics for available proteins sourced from the MR-Base NHGRI-EBI GWAS Catalog (https://gwas.mrcieu.ac.uk/) as mediators, which included 3,282 plasma proteins from 3,301 healthy participants in the INTERVAL study [12], 1,124 blood circulating proteins measured in 1,000 blood samples from the KORA study [13], and 83 proteins from 3,394 individuals in the IMPROVE study [14].

Identification of qualified genetic instrumental variables (IVs)

To identify qualified genetic instruments, we selected single-nucleotide polymorphisms (SNPs) based on stringent criteria: we initially extracted SNPs that achieved genome-wide significance ($P < 5E\text{-}08$) and then clumped them within a 1000 kb window to an LD

threshold of $R^2 < 0.1$, using the 1000 Genomes European Ancestry reference panel [15] to ensure genetic independence. To avoid the influence of weak instrument bias, we calculated the F statistic with formula: $F = R^2 \times (N - k - 1)/k \times (1 - R^2)$, and the genetic variation (R^2) with the formula $R^2 = 2 \times EAF \times (1 - EAF) \times \beta^2$, where N represents the sample size, k represents the number of IVs, EAF is the effect allele frequency, and β is the estimated effect size. SNPs with an F statistic of ≥10 were retained to minimize weak instrument bias. Finally, we harmonized the exposure and outcome datasets to ensure that the effect of each variant on both the exposure and outcome aligned with the same allele. We inferred positive-strand alleles and systematically excluded palindromic SNPs with ambiguous allele frequencies and any incompatible alleles.

Univariable MR (UVMR) and multivariable MR (MVMR)

In our study, we employed the R package "TwoSampleMR" (version: 0.5.8) for UVMR analyses to explore the relationship between circulating metabolites and DR outcomes. For single IV analyses, we utilized the Wald Ratio to estimate causal relationships. Under the assumption of valid IVs and no horizontal pleiotropy, we predominantly used the inverse-variance weighted (IVW) method as a robust MR approach to infer causality [16]. MVMR extends the standard MR framework by considering multiple potential risk factors within a single model [17]. To circumvent the suboptimal performance of traditional MVMR using standard linear regression in the presence of numerous risk factors, we employed MR-BMA. This Bayesian model averaging approach not only scales effectively to high-throughput experiments but also demonstrates robustness in detecting true causal risk factors, even when candidate factors are highly correlated [18]. The MVMR analyses were performed exclusively on causal circulating metabolites associated with the same DR outcome to ensure robust evaluation. Qualified IVs commonly related to causal circulating metabolites were extracted and processed using the same criteria as described previously ($P < 5E\text{-}08$, clumped at $R^2 < 0.1$ within a 1000 kb window, based on the 1000 Genomes European Ancestry reference panel). Within a Bayesian framework, MR-BMA calculates the marginal inclusion probability (MIP) and the model-averaged causal effect (MACE) for each risk factor. MIP is the sum of the posterior probabilities (PP) of all models that include the risk factor, indicating the likelihood that the risk factor is a causal determinant of disease risk. MACE provides a conservative estimate of the average direct causal effect of the risk factor on the outcome, derived through weighted averaging, with the weights determined by the posterior probabilities of the respective models.

Mediation MR analysis

We performed two-step MR analyses to explore RNA molecules that may mediate the link between circulating metabolites and DR outcomes. Initially, we applied UVMR to estimate the causal effect (β1) of circulating metabolites on each potential mediator. Subsequently, we also used UVMR to estimate the causal effect (β2) of each mediator on DR outcomes. If the results indicated that both β1 and β2 were significant, we used the "product of coefficients" method to calculate the mediation effect (β1 × β2) of circulating metabolites on DR outcomes through each mediator. We also calculated the direct effect of circulating metabolites on DR outcomes by excluding the mediator, which was derived by subtracting the mediation effect from the total effect. The standard errors for the mediation effects were calculated using the delta method formula: $SE_{mediation} = \sqrt{(\beta 1 \times SE1) + (\beta 2 \times SE2)}$. The z-score for the mediation effects was then calculated as: $Z = \text{mediation effect}\ (\beta 1 \times \beta 2)/SE_{mediation}$. Finally, the *P*-value for the mediation effects was calculated using the formula: $P = 2 \times \text{pnorm}(q = |Z|, lower.tail = FALSE)$. Negative mediation proportions were truncated at a minimum threshold of 0%, as this is the lowest threshold to determine a mediation proportion.

Sensitivity analysis

For the UVMR analyses, we conducted several sensitivity analyses to support the IVW estimates, including the weighted median, simple mode, weighted mode, and MR-Egger. The weighted median approach provides a reliable estimate of causality when at least 50% of the weight is derived from valid IVs [19]. The simple and weighted mode methods estimate the causal effect based on mode of the unweighted and IVW empirical density functions, respectively [20]. The MR-Egger regression method can detect directional horizontal pleiotropy and provide a corrected estimate [21]. The *P*-value < 0.05 for the MR-Egger intercept indicates the presence of directional pleiotropy. Cochran's Q statistic was calculated to evaluate heterogeneity [22].

Functional annotation for RNA mediators

To annotate the RNA mediators implicated in mediating the causal relationship between metabolites and DR outcomes, we conducted functional annotation to uncover their biological significance. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the R packages "clusterProfiler" and "org.Hs.eg.db" to identify pathways enriched with the identified RNA mediators. Additionally, we constructed a Protein-Protein

Interaction (PPI) network using data from the STRING database.

Statistical analysis

To address multiple testing, we employed the Benjamini-Hochberg false discovery rate (FDR) procedure. IVW estimates with $|\beta| > 0.1$, P-value < 0.05, and FDR < 0.05, supported by at least one sensitivity analysis, were considered robust evidence of causality. IVW estimates with a P-value < 0.05 but FDR ≥ 0.05 or lacking support from sensitivity analyses were considered suggestive of potential causality. To clearly interpret the causal effect (β), we utilized the odds ratio (OR), an intuitive indicator for assessing risk, to exhibit the potential impact of lipid levels on DR outcomes. All MR analyses were performed using R software (version 4.3.1) with the following R packages: "TwoSampleMR" (version 0.5.8), "MVMR" (version 0.4), and "MendelianRandomization" (version 0.10.0). Additionally, R packages "clusterProfiler" (version 4.10.1) and "org.Hs.eg.db" (version 3.16.0) were utilized for GO and KEGG pathway investigations. The code for the MR-BMA method was obtained from the GitHub repository (https://github.com/verena-zuber/demo_AMD) referenced in the literature [18].

Results

Two putative causal major lipids for DR outcomes with robust evidence

In the UVMR analyses of five major lipids, including triglycerides, apolipoprotein B, LDL cholesterol, HDL cholesterol, and apolipoprotein A-I, we observed significant causal relationships between nearly all these lipids and one or more of the four DR outcomes (P-value < 0.05, as shown in Fig. 1, Table S2). Specifically, elevated triglyceride levels increased the risks for DR (OR [95% CI] = 1.11 [1.07–1.16]; P = 7.57e-06) with robust causal evidence, background DR (OR [95% CI] = 1.07[1.01–1.14]; P = 0.03), and severe background DR (OR [95% CI] = 1.18 [1.04–1.33]; P = 0.02) with suggestive causal evidence. Higher HDL cholesterol levels decreased the risk for all four DR outcomes, with robust evidence of causality for background DR (OR [95% CI] = 0.88[0.83–0.93]; P = 2.66e-06) and the strongest protective effect for severe background DR (OR [95% CI] = 0.82[0.70–0.93]; P = 5.20e-04). Additionally, apolipoprotein A-I exhibited varying degrees of protective effects across the four DR outcomes, with suggestive evidence of causality. Both LDL cholesterol and apolipoprotein B exhibited slight protective effects against DR (OR = 0.92 and 0.95 respectively), LDL cholesterol showed suggestive causal effects on PDR (OR [95% CI] = 0.95[0.90–0.99]; P = 0.015). However, it is important to note that most results exhibited significant heterogeneity and pleiotropy (P < 0.05). For results demonstrating pleiotropy, the MR-Egger method was employed to

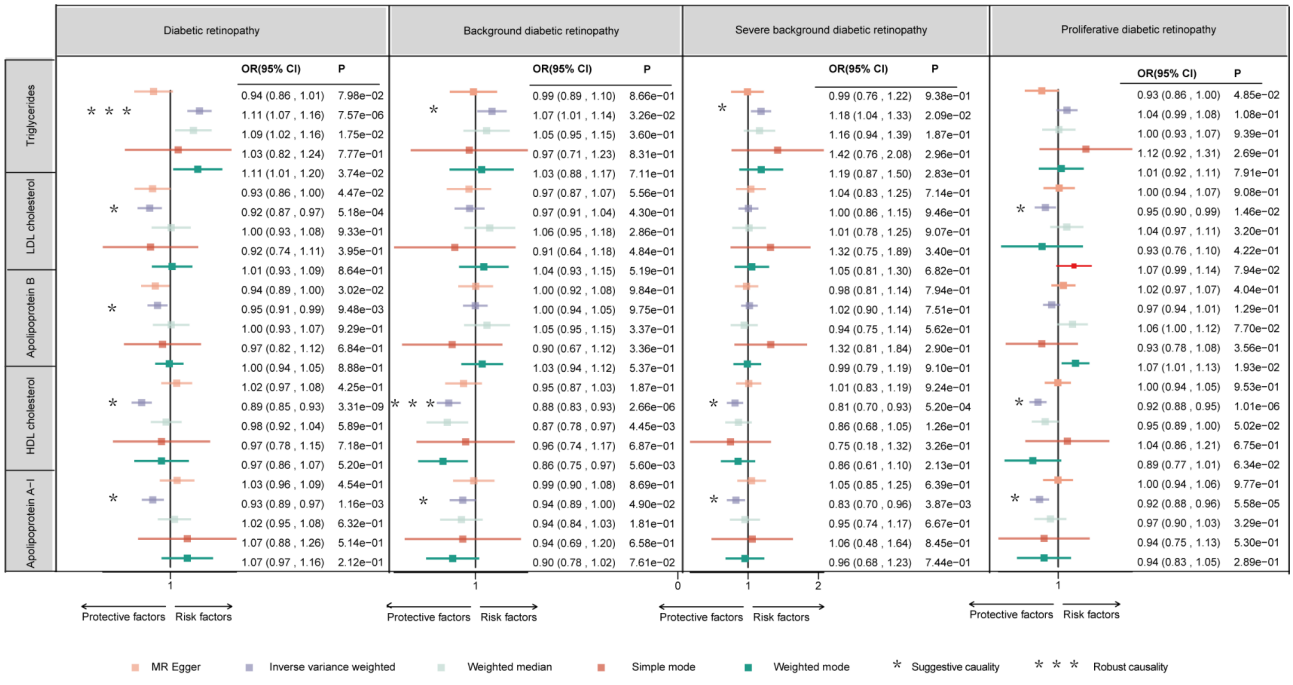

Fig. 1 Putative causal relationships between 5 major lipids and four DR outcomes. The forest plots delineated the causal impact of five major lipids on four diabetic retinopathy. (DR) outcomes through univariable Mendelian randomization analysis. An asterisk (*) indicates. suggestive causality, while three asterisks (***) denote robust causality

identify and exclude outlier SNPs with horizontal pleiotropy, ensuring the robustness of the conclusions (Table S3). For results indicating heterogeneity, the weighted median approach was utilized to facilitate causal inference (Table S4).

Eleven plasma proteins mediated the causal effect of two causal major lipids on DR outcomes

Our findings indicate that elevated triglyceride levels are associated with an increased risk of DR, whereas higher HDL cholesterol levels are linked to a reduced risk of background DR, substantiated by robust causal evidence. Mediation MR analysis identified ten plasma proteins (Fig. 2A) that significantly mediate the causal effect of triglycerides on DR (β = 0.12). Among these,

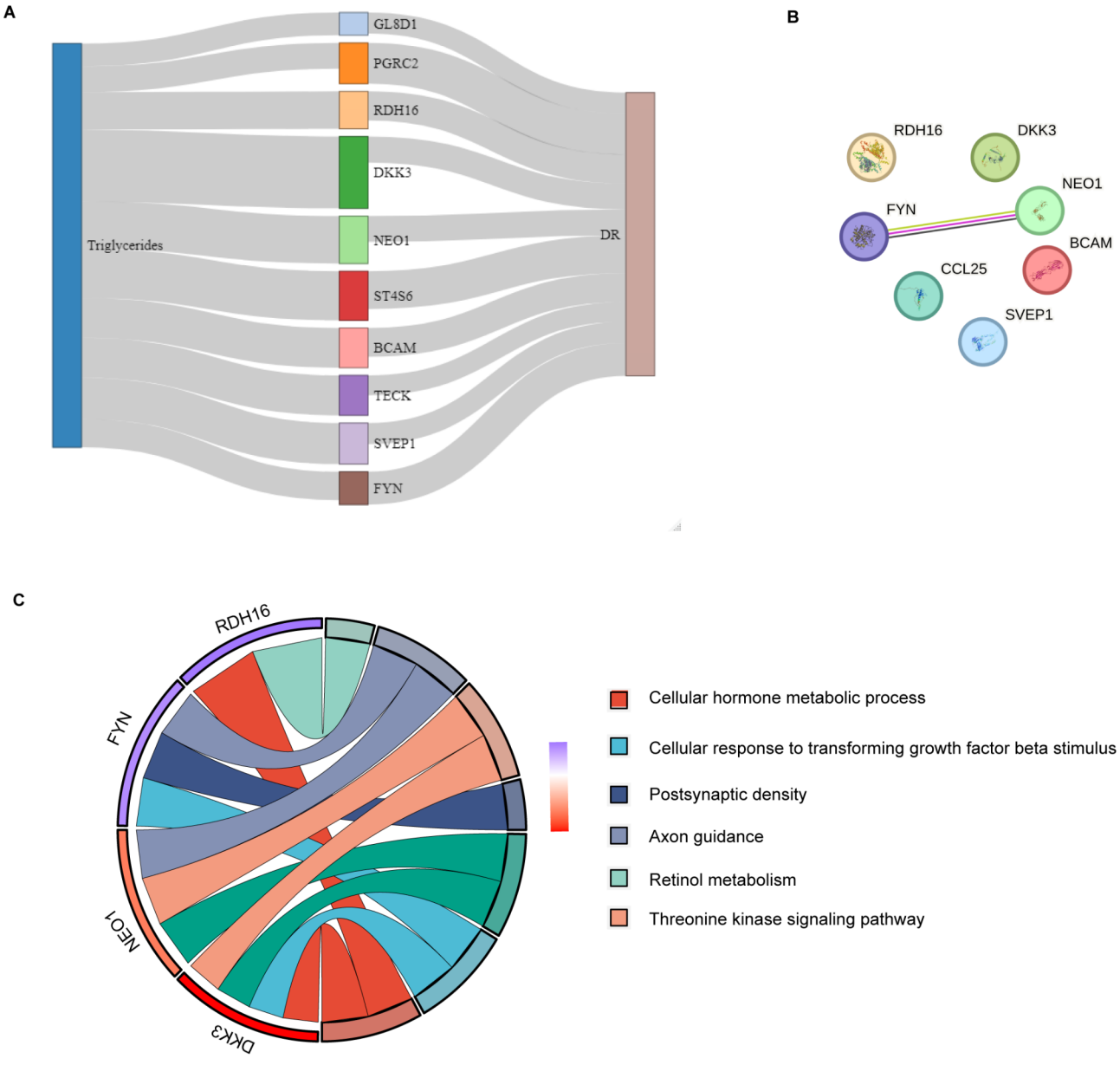

Fig. 2 Proteins mediating the causal effect of triglycerides on DR. The chord diagram illustrated that ten proteins significantly mediated the causal effect of triglycerides on diabetic retinopathy(DR). **A**. Protein-protein interaction (PPI) analysis revealed that among these ten proteins, only Proto-Oncogene, Src Family Tyrosine Kinase (FYN), and Neogenin 1 (NEO1) exhibited interactions. **C**. Enrichment analyses demonstrated that Retinol Dehydrogenase 16 (RDH16), Dickkopf-related protein 3 (DKK3), NEO1, and FYN were significantly implicated in various biological processes, including Cellular Hormone Metabolic Process (GO: BP), Cellular Response to Transforming Growth Factor Beta Stimulus (GO: BP), Threonine Kinase Signaling Pathway (GO: BP), Postsynaptic Density (GO: CC), Co-receptor Binding (GO: MF), Axon Guidance (KEGG), and Retinol Metabolism (KEGG) (Fig. 2C, adjusted $P < 0.05$)

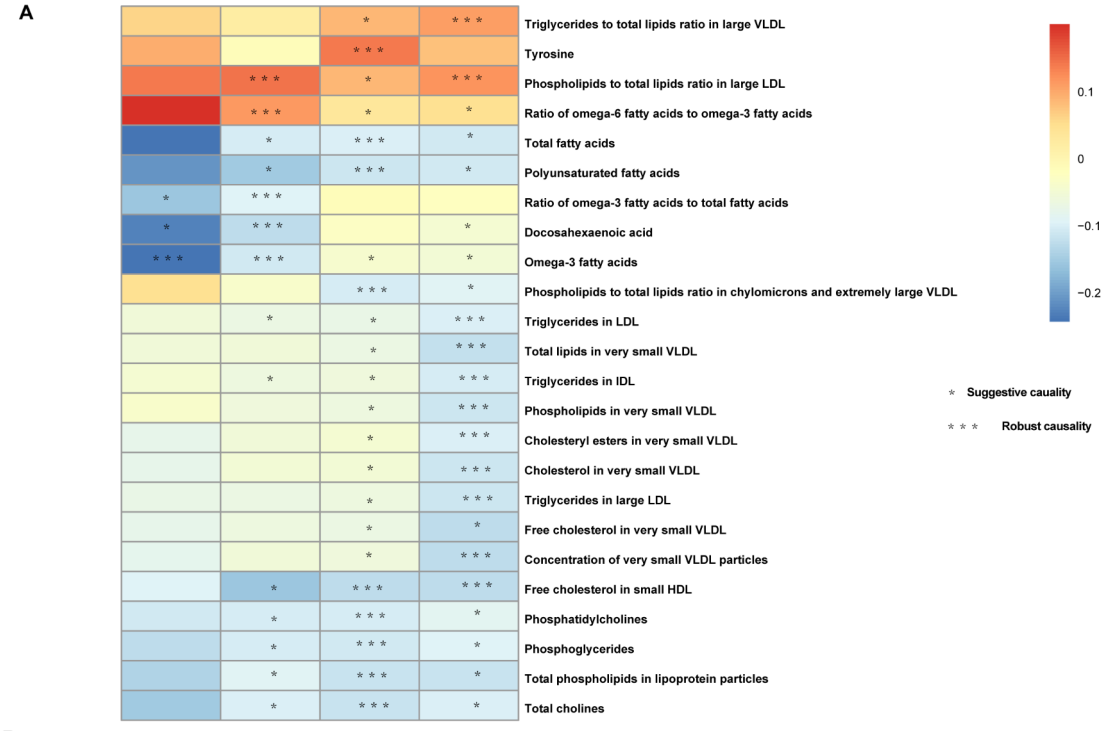

Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Putative causal circulating metabolites of four DR outcomes identified by univariable and multivariable Mendelian randomization analyses. **A**. The heatmap illustrated the effects of 26 causal circulating metabolites with robust evidence on four diabetic retinopathy (DR) outcomes identified through univariable Mendelian randomization (UVMR) analyses. These included 11 metabolites for overall DR, 5 for background DR, 1 for severe background DR, and 9 for proliferative DR (PDR). **B**. The forest plots depicted the results of multivariable Mendelian randomization (MVMR) analyses using the MR-BMA method. These analyses identified 8 out of the 11 causal metabolites for overall DR, 1 out of the 5 for background DR, and 2 out of the 9 for PDR, considering the combined impact of these metabolites

Dickkopf-related protein 3 (DKK3), Sulfotransferase Family 4 A, Memb.

-er 1 (ST4S6), and Neogenin 1 (NEO1) exhibit mediation effects consistent with the overall impact of triglycerides on DR (β = 0.03, 0.03, and 0.02 in mediation MR analyses, respectively). Additionally, PPI analysis revealed an interaction between Proto-Oncogene, Src Family Tyrosine Kinase (FYN) and NEO1 within the set of ten proteins (Fig. 2B). Enrichment analyses demonstrated that Retinol Dehydrogenase 16 (RDH16), DKK3, NEO1, and FYN were significantly involved in biological processes such as Cellular Hormone Metabolic Process (GO: BP), Cellular Response to Transforming Growth Factor Beta Stimulus (GO: BP), Threonine Kinase Signaling Pathway (GO: BP), Postsynaptic Density (GO: CC), Co-receptor Binding (GO: MF), Axon Guidance (KEGG), and Retinol Metabolism (KEGG) (Fig. 2C, adjusted $P < 0.05$). Furthermore, mediation MR analysis also revealed that the protein RPN1 significantly mediated the causal effect of HDL cholesterol on background DR (β = -0.10), with a mediation effect of -0.01 (Table S5).

Twenty-six putative causal metabolites for DR outcomes with robust evidence

In the UVMR analyses, eleven metabolites, including seven associated with very low-density lipoprotein (VLDL), three with low-density lipoprotein (LDL), and one with intermediate-density lipoprotein (IDL), demonstrated robust causal associations with DR risk (Fig. 3A and Table S2). Among these, the elevated triglycerides to total lipids ratio in large VLDL (OR [95% CI] = 1.11 [1.06–1.16]; $P = 3.71e-05$) and increased phospholipids to total lipids ratio in large LDL (OR [95% CI] = 1.12 [1.06–1.17]; $P = 2.66e-06$) were found to significantly heighten the DR risk. The remaining nine metabolites exhibited negative causal associations with DR risk. However, when considering the combined effect of these eleven metabolites on DR risk through MR-BMA analysis, only six VLDL-related metabolites and two LDL-related metabolites exhibited significant negative associations with DR. Specifically, cholesteryl esters in very small VLDL, cholesterol in very small VLDL, concentration of very small VLDL particles, and triglycerides in large LDL showed higher MIPs of 0.918, 0.814, 0.744, and 0.631, respectively, with FDRs < 0.05.

For background DR, UVMR analyses identified five metabolites with robust causal associations. Among these, two metabolites related to omega-3 fatty acids (OR = 0.89 and 0.91; $P = 6.87e-05$ and 8.97e-05) and docosahexaenoic acid (OR [95% CI] = 0.88 [0.82–0.94]; $P = 5.38e-05$) were observed to have significant negative associations with background DR risk. Conversely, an elevated ratio of omega-6 fatty acids to omega-3 fatty acids (OR [95% CI] = 1.12 [1.07–1.17]; $P = 2.40e-05$) and an increased phospholipids to total lipids ratio in large LDL (OR [95% CI] = 1.15 [1.09–1.22]; $P = 3.83e-05$) were associated with a heightened background DR risk. MR-BMA analysis, considering the combined impact of these five causal metabolites, found that only the phospholipids to total lipids ratio in large LDL had a significant effect on background DR risk with MIPs of 0.943 and FDR < 0.05(Fig. 3B). For severe background DR, UVMR analyses identified that only elevated omega-3 fatty acid levels were significantly associated with a reduced risk of severe background DR (OR [95% CI] = 0.78 [0.66–0.90]; $P = 7.00e-05$).

Regarding PDR, UVMR analyses identified nine metabolites with robust causal associations. Among these, increased tyrosine levels were significantly associated with an elevated risk of PDR (OR [95% CI] = 1.15 [1.08–1.21]; $P = 5.50e-05$). Additionally, two metabolites related to fatty acids (OR = 0.90 and 0.89; $P = 1.29e-04$ and 1.11e-05), two phospholipid-related metabolites (OR = 0.88 and 0.90; $P = 9.50e-05$ and 1.56e-05), two choline-related metabolites (OR = 0.89 and 0.90; $P = 2.28e-05$ and 1.59e-5), free cholesterol in small HDL (OR [95% CI] = 0.88 [0.81–0.94]; $P = 9.55e-05$), and phosphoglycerides (OR [95% CI] = 0.89 [0.84–0.94]; $P = 1.08e-05$) were observed to have significant negative associations with PDR. MR-BMA analysis, considering the combined impact of these nine causal metabolites, revealed that only elevated levels of total cholines and polyunsaturated fatty acids (PUFAs) were significantly associated with a reduced risk of PDR with MIPs of 0.164 and 0.184, respectively, and all FDR < 0.05(Fig. 3B).

Thirteen plasma proteins mediated the effect of causal metabolites on DR outcomes

Although our study initially identified 11 metabolites with robust causal associations with DR risk, we subsequently validated 8 of these metabolites through MVMR analysis. The proteome-wide mediation MR analysis identified 10 plasma proteins that mediate the causal effects of 7 of these metabolites on DR (Table 1 and Table S5). For the 5

Table 1 Summary of significant plasma proteins from mediation MR analysis of causal metabolites on DR outcomes

Exposure	Mediate proteins	Outcome	Total Effect	Effect of Exposure on Mediator	Effect of Mediator on Outcome	Mediation Effect	Proportion Mediated	Direct Effect	*P*-value
Triglycerides in large LDL	BCAM	DR	-0.119	-0.194	0.122	-0.024	0.198	-0.096	1.39E-02
Triglycerides in large LDL	TMEM2	DR	-0.119	-0.122	-0.123	0.015	-0.126	-0.134	3.47E-02
Triglycerides in large LDL	SVEP1	DR	-0.119	-0.117	0.087	-0.01	0.085	-0.109	4.26E-02
Concentration of very small VLDL particles	ST4S6	DR	-0.136	-0.188	-0.16	0.03	-0.221	-0.166	6.46E-03
Concentration of very small VLDL particles	CLIC5	DR	-0.136	-0.22	0.287	-0.063	0.464	-0.073	3.07E-03
Concentration of very small VLDL particles	BCAM	DR	-0.136	-0.129	0.122	-0.016	0.116	-0.12	4.24E-02
Concentration of very small VLDL particles	DKK3	DR	-0.136	-0.122	-0.11	0.013	-0.099	-0.15	4.80E-02
Phospholipids in very small VLDL	N- terminal pro-BNP	DR	-0.126	-0.128	-0.123	0.016	-0.125	-0.142	3.02E-02
Phospholipids in very small VLDL	ST4S6	DR	-0.126	-0.17	-0.16	0.027	-0.215	-0.153	8.58E-03
Phospholipids in very small VLDL	CLIC5	DR	-0.126	-0.207	0.287	-0.059	0.469	-0.067	4.96E-03
Phospholipids in very small VLDL	DKK3	DR	-0.126	-0.138	-0.11	0.015	-0.121	-0.142	3.20E-02
Triglycerides in LDL	BCAM	DR	-0.108	-0.186	0.122	-0.023	0.211	-0.085	1.87E-02
Triglycerides in LDL	SVEP1	DR	-0.108	-0.144	0.089	-0.013	0.119	-0.095	3.80E-02
Free cholesterol in very small VLDL	BCAM	DR	-0.147	-0.13	0.122	-0.016	0.108	-0.131	4.49E-02
Free cholesterol in very small VLDL	PGRC2	DR	-0.147	-0.093	-0.173	0.016	-0.11	-0.163	1.76E-02
Total lipids in very small VLDL	ST4S6	DR	-0.134	-0.17	-0.146	0.025	-0.185	-0.158	1.01E-02
Total lipids in very small VLDL	THTR	DR	-0.134	-0.08	-0.085	0.007	-0.051	-0.14	4.59E-02
Total lipids in very small VLDL	BCAM	DR	-0.134	-0.161	0.122	-0.02	0.147	-0.114	2.27E-02
Total lipids in very small VLDL	CLIC5	DR	-0.134	-0.214	0.287	-0.061	0.458	-0.072	4.06E-03
Total lipids in very small VLDL	SIRPG	DR	-0.134	0.133	0.212	0.028	-0.211	-0.162	2.91E-02
Total lipids in very small VLDL	DKK3	DR	-0.134	-0.129	-0.11	0.014	-0.106	-0.148	3.79E-02
Cholesteryl esters in very small VLDL	ST4S6	DR	-0.109	-0.12	-0.146	0.018	-0.162	-0.126	3.46E-02
Cholesteryl esters in very small VLDL	SIRPG	DR	-0.109	0.14	0.212	0.03	-0.273	-0.138	2.38E-02
Cholesteryl esters in very small VLDL	CLIC5	DR	-0.109	-0.196	0.287	-0.056	0.516	-0.053	7.50E-03
Phospholipids to total lipids ratio in large LDL	NAD(P)H dehydrogenase	Background DR	0.135	0.099	0.14	0.014	0.103	0.121	1.68E-02
Polyunsaturated fatty acids	RFNG	PDR	-0.118	0.382	0.194	0.074	-0.63	-0.192	1.80E-03
Polyunsaturated fatty acids	PDE4D	PDR	-0.118	0.107	-0.234	-0.025	0.212	-0.093	3.55E-02
Total cholines	RFNG	PDR	-0.107	0.312	0.194	0.061	-0.564	-0.168	8.24E-03

very small VLDL-related metabolites, proteins Chloride Intracellular Channel 5 (CLIC5) and BCAM (Basal Cell Adhesion Molecule) significantly potentiated the protective effects of very small VLDL-related metabolites on DR (mediation β = -0.06 and – 0.02, total effect β = -0.11 to -0.15). Conversely, proteins (Signal Regulatory Protein Gamma) SIRPG, ST4S6, PGRC2 (Progestin and AdipoQ Receptor Family Member 2), DKK3, N-terminal pro-Brain Natriuretic Peptide (N-terminal pro-BNP), and (Thiamine Transporter) THTR significantly attenuated the protective effects of VLDL-related metabolites on DR (mediation β > 0). For the 2 LDL-related metabolites, proteins BCAM and Sushi, Von Willebrand Factor Type A, EGF And Pentraxin Domain Containing 1 (SVEP1) significantly augmented the protective effects of LDL-related metabolites on DR (mediation β = -0.02 and – 0.01, total effect β = -0.108 and – 0.119, respectively).

Additionally, for background DR, proteome-wide mediation MR analysis revealed that NAD(P)H dehydrogenase mediates the risk effect of the phospholipid to total lipid ratio in large LDL on background DR (total effect β = 0.14, mediation effect = 0.01). For PDR, protein RFNG O-Fucosylpeptide 3-Beta-N-Acetylglucosaminyltransferase (RFNG) mitigates the protective effects of PUFAs and total cholines on PDR (mediation β = 0.07 and 0.06, total effect β = -0.12 and – 0.11, respectively), whereas protein PDE4D positively mediates the protective effects of PUFAs on PDR (mediation β = -0.03, total effect β = -0.12).

Discussion

Among the major lipids analyzed, triglycerides emerged as the most significant risk factor, potentially contributing to disease risk through mediation by the proteins DKK3, ST4S6, and NEO1.Conversely, HDL cholesterol was identified as the most potent protective factor, potentially reducing the risk of background DR through mediation by the protein RPN1. Among the causal circulating metabolites, cholesteryl esters in very small VLDL exhibited the strongest protective effect against DR, with their influence mediated by the plasma proteins CLIC5 and BCAM.For background DR, the phospholipid-to-total lipid ratio in large LDL was identified as the most plausible causal metabolite, mediated by the plasma protein NAD(P)H dehydrogenase. Additionally, RFNG and PDE4D positively mediate the protective effects of PUFAs on PDR.

Extensive evidence indicates that inadequate control of triglyceride levels is associated with the onset and progression of DR, whereas elevated HDL-C levels and the use of lipid-lowering medications significantly diminish the risk of DR [23–25]. Our results also revealed that the elevated triglyceride levels can distinctly increased DR risk, especially, we demonstrated that reduced HDL-C levels can increased background DR risk with robust causal evidence. Proteome-wide mediation MR analyses identified eleven plasma proteins that mediate the causal effects of triglyceride and HDL-C on DR outcomes, such as DKK3, ST4S6, NEO1, RDH16 and so on (Fig. 2A). DKK3, a crucial member of the DKK family and an important modulator of Wnt signaling, is synthesized and secreted by Muller cells. A study involving 44 eyes from 39 patients with diabetic macular edema (DME) and 27 eyes from 27 controls identified significantly elevated levels of DKK3 in the aqueous humor of DME patients and in human Müller cells. This research suggests that excessive activation of Wnt signaling, mediated by elevated DKK3 levels, may contribute to neovascularization and the progression of DR [26]. Conversely, RDH16 is integral to retinal health due to its role in efficiently producing retinal reductase, which supports retinol metabolism [27, 28]. Our research also indicated that RDH16 negatively mediated the risk impact of triglycerides on DR.

Based on the analyses from both UVMR and MVMR, we have identified eight lipoprotein subclasses (including six very small VLDL particles and two LDL particles) that are protectively associated with DR. Very small VLDL particles have also reported to be inversely associated with incident diabetes [29] and age-related macular degeneration (AMD) [30]. Furthermore, a population-based study involving Chinese, Malay, and Indian adults used logistic regression to find that certain very small VLDL particles, consistent with our findings, such as cholesteryl esters in very small VLDL and LDL particles (including total lipids in large LDL), were protectively associated with moderate or more severe DR [31]. Additionally, the protein CLIC5 was found to significantly mediate the protective causal effect of very small VLDL particles on DR. While CLIC5 has also been reported to be significantly downregulated in glomerular tissues of diabetic nephropathy patients [32]. This suggests that the protective effect of very small VLDL particles on DR may, in part, be mediated through CLIC5. Our study elucidated that NAD(P)H dehydrogenase mediated the risk effect of the phospholipid-to-total lipid ratio in large LDL on background DR. NAD(P)H dehydrogenase is pivotal in modulating cellular redox balance and energy metabolism. Previous studies have shown that diabetic rats display elevated concentrations of free NAD(P)H, reflecting increased glycolytic activity, along with higher levels of bound NAD(P)H, suggesting enhanced oxidative phosphorylation in their retinas [33]. These observations suggest that alterations in NAD(P)H dynamics, driven by modifications in lipid profiles, may exacerbate oxidative stress and metabolic dysregulation.

Our study further revealed that PUFAs and total choline exhibited protective effects against PDR. PUFAs were generally considered beneficial [34, 35]. Notably, two clinical studies conducted in Europe have identified an inverse relationship between omega-6 PUFAs and the incidence of DR [36, 37]. Additionally, elevated phosphatidylcholine levels are associated with reduced risks of diabetes and cardiovascular diseases [38]. Moreover, our findings indicate that RFNG positively mediate the protective effects of PUFAs and total cholines on PDR. RFNG enhances NOTCH1 activity by modifying O-fucose residues on specific EGF-like domains, thereby promoting NOTCH1 activation through DLL1 and JAG1, which may contribute to neurogenesis [39]. The interaction between PUFAs, choline, and RFNG underscores a complex network of metabolic and signaling pathways that collectively influence retinal health. These insights could inform future therapeutic strategies aimed at leveraging these protective factors to prevent or mitigate the progression of DR.

Limitations

Despite the significant findings, our study has several limitations. First, although MR helps to infer causality by minimizing confounding and reverse causation, it relies on the assumption of no pleiotropy, meaning the genetic variants used as instruments should not affect the outcome through pathways other than the exposure of interest. Violations of this pleiotropy assumption could bias the results. Second, our analysis was based on data from European participants, which may limit the generalizability of our findings to other ethnic and demographic

groups. Third, while we identified several proteins that potentially mediate the effects of lipids on DR, the exact biological mechanisms remain to be elucidated through experimental studies. Additionally, the use of plasma lipid measurements might not fully reflect lipid metabolism within retinal tissues, and tissue-specific studies are needed to confirm our findings. Furthermore, our study relies on summary-level GWAS data, which, while enabling large-scale causal inference, comes with inherent limitations. The use of summary statistics precludes individual-level data analyses, limiting our ability to assess potential confounding factors and effect modifications at a finer scale. Independent cohort validation is essential to confirm the robustness and replicability of our findings across diverse populations. Finally, the complexity of lipid metabolism and its interaction with various metabolic pathways necessitates further investigation to fully understand the causal relationships identified in this study. Future research should focus on longitudinal and tissue-specific analyses to validate and extend our findings.

Conclusion

In conclusion, our metabolome-wide MR analysis has elucidated the complex relationships between lipid profiles and DR. We identified triglycerides as a significant risk factor for DR, mediated by the proteins DKK3, ST4S6, and NEO1, while HDL-C emerged as a potent protective factor, potentially reducing the risk of background DR through RPN1 mediation. Cholesteryl esters in very small VLDL exhibited the strongest protective effect against DR, mediated by CLIC5 and BCAM, while the phospholipid-to-total lipid ratio in large LDL was identified as a key causal metabolite for background DR, with its effects mediated by NAD(P)H dehydrogenase. Furthermore, the protective effects of PUFAs and total cholines on PDR were positively mediated by RFNG. These findings provide valuable insights into the metabolic pathways and potential therapeutic targets for DR, highlighting the importance of lipid metabolism in the pathogenesis of this condition. Future research should focus on validating these results in diverse populations and exploring the underlying mechanisms through experimental studies.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13098-025-01701-z.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Author contributions

Jiawei Wang: conceptualization; data curation; formal analysis and writing original draft; Jing Su: methodology; software; and validation; Jingxue Ma and Danyan Liu: funding acquisition; investigation; project administration; resources; supervision; visualization; and revision. All authors reviewed the manuscript.

Funding

The work was supported by Medical Science Research Project of Hebei Province [20221109], and Hospital-level Foundation of the Second Hospital of Hebei Medical University [2HC202020].

Data availability

The datasets employed in this study are publicly accessible summary datasets. They can be found in the cited publications, within the IEU OpenGWAS Project repository ([https://gwas.mrcieu.ac.uk/], or in the FinnGen repository ([https://storage.googleapis.com/
finngen-public-data-r9/summary_stats/].

Declarations

Ethics approval and consent to participate

The summary datasets utilized in this study are publicly accessible and available through the referenced publications. All original GWAS investigations incorporated in this research were conducted with the approval of their respective ethics committees, and informed consent was obtained from all participants involved in these studies.

Competing interests

The authors declare no competing interests.

Received: 26 January 2025 / Accepted: 12 April 2025
Published online: 26 April 2025

References

1. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and Meta-analysis. Ophthalmology. 2021;128:1580–91.
2. Leasher JL, Bourne RR, Flaxman SR, et al. Global estimates on the number of people blind or visually impaired by diabetic retinopathy: A Meta-analysis from 1990 to 2010. Diabetes Care. 2016;39:1643–9.
3. Global regional, national burden of diabetes. From 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the global burden of disease study 2021. Lancet. 2023;402:203–34.
4. Yahyavi SK, Snorgaard O, Knop FK, et al. Prediabetes defined by first measured HbA(1c) predicts higher cardiovascular risk compared with HbA(1c) in the diabetes range: A cohort study of nationwide registries. Diabetes Care. 2021;44:2767–74.
5. Chapman MJ, Sposito AC. Hypertension and dyslipidaemia in obesity and insulin resistance: pathophysiology, impact on atherosclerotic disease and pharmacotherapy. Pharmacol Ther. 2008;117:354–73.
6. Liu W, Yang C, Lei F, et al. Major lipids and lipoprotein levels and risk of blood pressure elevation: a Mendelian randomisation study. EBioMedicine. 2024;100:104964.
7. Li B, Zhao X, Xie W, et al. Causal association of Circulating metabolites with diabetic retinopathy: a bidirectional Mendelian randomization analysis. Front Endocrinol (Lausanne). 2024;15:1359502.
8. Hagenbeek FA, Pool R, van Dongen J, et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. Nat Commun. 2020;11:39.
9. Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between Circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. PLoS Med. 2020;17:e1003062.
10. Borges MC, Haycock PC, Zheng J, et al. Role of Circulating polyunsaturated fatty acids on cardiovascular diseases risk: analysis using Mendelian randomization and fatty acid genetic association data from over 114,000 UK biobank participants. BMC Med. 2022;20:210.

11. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613:508–18.
12. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558:73–9.
13. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017;8:14357.
14. Folkersen L, Fauman E, Sabater-Lleal M, et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. PLoS Genet. 2017;13:e1006706.
15. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491:56–65.
16. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37:658–65.
17. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2019;48:713–27.
18. Zuber V, Colijn JM, Klaver C, Burgess S. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. Nat Commun. 2020;11:29.
19. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol. 2016;40:304–14.
20. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal Pleiotropy assumption. Int J Epidemiol. 2017;46:1985–98.
21. Bowden J, Del Greco MF, Minelli C, et al. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol. 2016;45:1961–74.
22. Bowden J, Del Greco MF, Minelli C, et al. A framework for the investigation of Pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017;36:1783–802.
23. Sen K, Misra A, Kumar A, Pandey RM. Simvastatin retards progression of retinopathy in diabetic patients with hypercholesterolemia. Diabetes Res Clin Pract. 2002;56:1–11.
24. Chew EY, Ambrosius WT, Davis MD, et al. Effects of medical therapies on retinopathy progression in type 2 diabetes. N Engl J Med. 2010;363:233–44.
25. Keech AC, Mitchell P, Summanen PA, et al. Effect of Fenofibrate on the need for laser treatment for diabetic retinopathy (FIELD study): a randomised controlled trial. Lancet. 2007;370:1687–97.
26. Ji B, Lim D, Kim J, et al. Increased levels of Dickkopf 3 in the aqueous humor of patients with diabetic macular edema. Invest Ophthalmol Vis Sci. 2016;57:2296–304.
27. Parés X, Farrés J, Kedishvili N, Duester G. Medium- and short-chain dehydrogenase/reductase gene and protein families: Medium-chain and short-chain dehydrogenases/reductases in retinoid metabolism. Cell Mol Life Sci. 2008;65:3936–49.
28. Pavez Loriè E, Li H, Vahlquist A, Törmä H. The involvement of cytochrome p450 (CYP) 26 in the retinoic acid metabolism of human epidermal keratinocytes. Biochim Biophys Acta. 2009;1791:220–8.
29. Dugani SB, Akinkuolie AO, Paynter N, et al. Association of lipoproteins, insulin resistance, and Rosuvastatin with incident type 2 diabetes mellitus: secondary analysis of a randomized clinical trial. JAMA Cardiol. 2016;1:136–45.
30. Julian TH, Cooper-Knock J, MacGregor S, et al. Phenome-wide Mendelian randomisation analysis identifies causal factors for age-related macular degeneration. Elife. 2023;12:e82546.
31. Quek D, He F, Sultana R, et al. Novel serum and urinary metabolites associated with diabetic retinopathy in three Asian cohorts. Metabolites. 2021;11:614.
32. Fan C, Gao Y, Sun Y. Integrated multiple-microarray analysis and Mendelian randomization to identify novel targets involved in diabetic nephropathy. Front Endocrinol (Lausanne). 2023;14:1191768.
33. Su E, Kesavamoorthy N, Junge JA, et al. Comparison of retinal metabolic activity and structural development between rd10 mice and normal mice using multiphoton fluorescence lifetime imaging microscopy. Curr Issues Mol Biol. 2024;46:612–20.
34. Vaidya H, Cheema SK. Arachidonic acid has a dominant effect to regulate lipogenic genes in 3T3-L1 adipocytes compared to omega-3 fatty acids. Food Nutr Res. 2015;59:25866.
35. Marion-Letellier R, Savoye G, Ghosh S. Polyunsaturated fatty acids and inflammation. IUBMB Life. 2015;67:659–67.
36. Houtsmuller AJ, Zahn KJ, Henkes HE. Unsaturated fats and progression of diabetic retinopathy. Doc Ophthalmol. 1980;48:363–71.
37. Howard-Williams J, Patel P, Jelfs R, et al. Polyunsaturated fatty acids and diabetic retinopathy. Br J Ophthalmol. 1985;69:15–8.
38. Morze J, Wittenbecher C, Schwingshackl L, et al. Metabolomics and type 2 diabetes risk: an updated systematic review and Meta-analysis of prospective cohort studies. Diabetes Care. 2022;45:1013–24.
39. Kakuda S, Haltiwanger RS. Deciphering the Fringe-Mediated Notch code: identification of activating and inhibiting sites allowing discrimination between ligands. Dev Cell. 2017;40:193–201.

Publisher's note