# RESEARCH

# **Open Access**



# Integration of multi-omics transcriptome-wide analysis for the identification of novel therapeutic drug targets in diabetic retinopathy

Guoguo Yi<sup>1,2,7†</sup>, Zhengran Li<sup>2,3†</sup>, Yuxin Sun<sup>2,3†</sup>, Xinyu Ma<sup>2</sup>, Zijin Wang<sup>3</sup>, Jinken Chen<sup>5</sup>, Dong Cai<sup>6</sup>, Ziran Zhang<sup>3</sup>, Zejun Chen<sup>3</sup>, Fanye Wu<sup>3</sup>, Mingzhe Cao<sup>4</sup> and Min Fu<sup>2\*</sup>

# Abstract

**Background** Diabetic retinopathy (DR) is the most important complication of Type 2 Diabetes (T2D) in eyes. Despite its prevalence, the early detection and management of DR continue to pose considerable challenges. Our research aims to elucidate potent drug targets that could facilitate the identification of DR and propel advancements in its therapeutic strategies.

**Methods** A broad multi-omics exploration of DR was presented to decipher the drug targets of DR and proliferative diabetic retinopathy (PDR). Transcriptome-Wide Association Studies (TWAS), fine-mapping and conditional analysis were applied to unearth potential tissue-specific gene associations with DR. Summary Data-based Mendelian Randomization (SMR) provided secondary analysis of high confidence genes. Cis-instrument of druggable genes were extracted from the eQTLGen Consortium and PsychENCODE, facilitating drug-target MR supported by colocalization analysis. Phenome-Wide Association Studies (PheWAS) was conducted on the high confidence genes. Metabolomic and immunomic MR-profiling further augmented our research as complement.

**Results** TWAS identified multiple robust genetic loci in both DR and PDR (WFS1, RPS26, and SRPK1) through genetic associations across different tissues. Meanwhile, we have delineated both the commonalities and discrepancies between DR and PDR at the transcriptomic level, represented by DCLRE1B as the hub gene that DR progressed into PDR. SMR revealed 92 key DR-related genes and 55 PDR-related genes. HLA-DQ family genes have a frequent occurrence, while RPS26, WFS1 and SRPK1 were validated as the genetic network's linchpins. Drug-target MR casted ERBB3 and SRPK1 as candidate effector genes for DR and PDR susceptibility. In addition, metabolomics and immunomics analyses also revealed multifaceted pathogenic factors for DR.

**Conclusions** Our research offers targeted therapeutic insights for early-stage DR and facilitates multi-omic comparisons of it and PDR.

 $^{\dagger}\mbox{Guoguo}$  Yi, Zhengran Li and Yuxin Sun have contributed equally to the research.

\*Correspondence: Min Fu min\_fu1212@163.com Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

# Highlights

- Our research was augmented by an array of multi-omics methodologies, notably Transcriptome-Wide Association Studies (TWAS), Summary Mendelian Randomization (SMR), and Drug-Target Mendelian Randomization (MR). These advanced techniques facilitated a thorough investigation into the genetic foundations of diabetic retinopathy (DR). We employed stringent correction procedures to guarantee the integrity of our findings.
- Our study adopted a comprehensive strategy, initiating with an examination of various systemic tissues. We
  meticulously identified pivotal genes linked to DR. By probing into genomic, metabolomic, and immunohistological dimensions, we revealed determinants implicated in DR progression.
- Our research uncovered a series of unique loci associated with DR, notably including RPS26, WFS1, and SRPK1. Meanwhile, we proposed DCLRE1B as an important gene affecting DR exacerbated to severe PDR status, ERBB3, SRPK1 as common genes that affects both sides. The molecules in the results of DR play an important role in the early screening and treatment of DR. What's more, the difference between PDR and DR can be applied in curbing the progression of DR and help control the precious rescue and cure time window.

**Keywords** Diabetic retinopathy (DR), Mendelian randomization (MR), Transcriptome-wide association studies (TWAS), Summary data-based Mendelian randomisation (SMR), Drug-target MR, Multi-omics, Metabolomics, Immunomics

# Introduction

Diabetic retinopathy (DR) stands as a significant microvascular complication inherent in both type 1 and type 2 diabetes [1]. Its prominence as a leading cause of vision loss among the working-age population underscores its critical status in medical research and intervention [2]. Traditional treatments for DR encompass a spectrum of interventions, ranging from surgical modalities such as laser therapy, intravitreal injections, and vitrectomy to non-surgical approaches targeting the management of hyperglycemia, hypertension, and hyperlipidemia [2]. However, current surgical interventions for DR are typically reserved for advanced stages and frequently fall short in preserving vision [3]. Concurrently, pharmacological therapies have demonstrated limited efficacy, leaving patients at considerable risk of irreversible blindness [2]. It is imperative to shift focus towards early pathogenic stages to optimize patient treatment outcomes, a pursuit of paramount clinical significance as it directly impacts vision preservation and mitigates the societal health burden associated with DR. In our endeavor, a thorough exploration of pharmacological interventions and the identification of novel therapeutic targets are indispensable. By embracing a holistic approach, we can advance with more effective strategies to combat this sight-threatening condition.

DR is recognized as a systemic ailment characterized by intricate and multifaceted pathogenic pathways. The etiology of DR encompasses various proposed hypotheses, including the activation of NLRP3 inflammatory vesicles, neutrophil extracellular trap-mediated damage, circRNA induction, and thickening of the retinal capillary basement membrane [4]. Distinctive retinal alterations have been identified as potential discriminators between proliferative diabetic retinopathy (PDR) and DR, with the vitreous also showing promise as a differentiating factor [5–7]. Despite the informative nature of studies elucidating the pathogenic mechanisms underlying DR, definitive conclusions remain elusive. Leveraging genetic correlations with DR offers a promising avenue for pinpointing single nucleotide polymorphisms (SNPs), elucidating pivotal genes at the transcriptomic level, identifying target loci, and uncovering DR-related biomarkers.

Our comprehensive multi-omic investigation delved into both DR and its advanced form, PDR, commencing with genome-wide association data. Through Transcriptome-Wide Association Studies (TWAS), conditional analyses, and permutation testing, we unveiled transcriptomic connections to DR/PDR. The identified genes underwent annotation using the Gene Ontology database. Leveraging Summary Data-Based Mendelian Randomization (SMR) with expression quantitative trait loci (eQTL)-derived SNPs, we employed the HEIDI test for validation. Druggable genes sourced from drug databases underwent MR analysis, targeting DR/PDR while adjusting for co-localization and false discovery rate (FDR). Additionally, our study encompassed Phenome-Wide Association Studies (PheWAS) and metabolomics-MR analyses to glean pharmacogenetic insights. An immunohistology-MR scan of 731 immune markers further elucidated key molecules. Despite significant advancements in existing therapies, such as anti-VEGF agents and laser treatments, substantial limitations remain. These treatments often do not adequately address the multifactorial nature of diabetic retinopathy, leading to the urgent need for more effective options that can comprehensively manage the condition [8]. This study identifies critical genes, such as RPS26, WFS1, and SRPK1, as novel therapeutic

targets, thereby paving the way for improved treatment strategies and advancing the field of diabetic retinopathy research.

# Methods

The overview of the study could be viewed in Fig. 1.

# **DR traits**

In our multi-omic analysis, we identified SNPs associated with DR using summary statistics from the FinnGen R9 dataset, which includes 319,046 individuals and 10,413 DR cases [9]. FinnGen provides a valuable resource for exploring genetic variants tied to disease progression, particularly within isolated populations. To further enrich the analysis, we integrated data on PDR from the FinnGen R6 dataset, which consists of 253,168 individuals, including 10,860 PDR cases. This combined approach enhances the representativeness and robustness of our findings [9].

# Gene expression weights for transcriptomic imputation

To explore the relationship between SNP transcription and diabetic retinopathy outcomes in a tissue-nonspecific manner, we utilized sparse canonical correlation analysis (sCCA) across multiple tissues to obtain gene expression weights [10]. For histologically localized transcriptomic insights into diabetic retinopathy, we focused on three distinct tissues—pancreas, kidney, and whole blood—using transcriptomic expression data from the GTEx V8 database [11]. This approach allowed us to assess tissue-specific gene expression profiles relevant to the progression of diabetic retinopathy.



Fig. 1 Study overview. An overview of the data sources, analysis process, methodology and broad results of this study. TWAS transcriptome-wide association study, SNP single nucleotide polymorphism, eQTL expression quantitative trait loci, FDR false discovery rate, PP.H4 posterior probability of H4, GTEx Genotype-Tissue Expression Project, PheWAS phenome-wide association study, IVW inverse variance weighted, DR diabetic retinopathy, MR-PRESSO Mendelian Randomization Pleiotropy RESidual Sum and Outlier test, MR Mendelian randomization

### Transcriptomic imputation

We applied a transcriptome-wide association study (TWAS) approach [12], which mediates the relationship between genetic variation and transcriptional regulation, transforming associations between individual genetic variants and phenotypes into gene or transcriptlevel insights. To predict tissue-specific gene expression, we used the FUSION pipeline, leveraging our genomewide association study (GWAS) summary statistics [12]. Unlike focusing solely on individual SNPs, FUSION allows for gene-level analysis, offering deeper insight into gene function and uncovering additional associations. The pipeline selects the most appropriate algorithm (such as susie, lasso, enet, or topl) to generate significant findings. We identified key genes by applying Bonferroni correction and used false discovery rate (FDR) correction for supplementary analysis. Specifically, we considered TWAS P < 0.05/ feature number each tissue after Bonferroni correction (Pancreas of 5734, Whole blood of 7938, Kidney cortex of 1203) and P < 1.32 e-06 of sCCA.

# Colocalization, permutation testing, and conditional analyses

Colocalization analysis for TWAS-significant genes (with a significance threshold of P < 0.05) was conducted using the coloc R package (version 5.1.0.1) within the FUSION framework [13]. This Bayesian-based method estimates the posterior probability (PP) of five models related to associations between GWAS and TWAS results. Specifically, we assessed the probability that our TWAS associations reflected either: (1) PP.H3: Linkage between distinct causal SNPs. (2) PP.H4: A single causal SNP. By comparing PP.H3 and PP.H4, we discerned whether a genuine causal relationship exists between traits and DR. We defined positive genes as those with PP.H4 exceeding 0.9. To validate our findings, we employed permutation testing and considered a P-value < 0.05 as a positive indicator [14]. Additionally, we utilized conditional analysis to identify co-expression sites and distinguish between independent and conditional features [12]. Conditioned/ marginal features were significantly associated with DR only in the unadjusted model, where their associations depended entirely on the expression of other proximate features. After correction, independently/jointly significant features remained associated with the phenotype at a nominal significance level (p < 0.05). Our conditional analysis followed the procedures outlined on the FUSION webpage (http://gusevlab.org/projects/fusion/).

# Fine-mapping of TWAS associations and high confidence features

We employed FOCUS to identify genes with causal associations to DR. FOCUS is a TWAS fine-mapping

method that, similar to the statistical fine-mapping of GWAS results, estimates the posterior inclusion probability (PIP) of each feature being causally linked within the association region. A PIP value greater than 0.5 suggests that the locus is more likely to be causally associated with DR than any other locus in the region. FOCUS accommodates multiple causal SNPs and genes while integrating gene effect sizes using conjugate priors. In our analysis, we defined high-confidence genes as those with PIP > 0.8, a notable TWAS P-value, and a significant conditional analysis P-value [15].

# Functional annotations of significant loci

To further explore the functions of key loci and genes, we annotated the results of transcriptome analyses on Gene Ontology (GO) databases accordingly to interpret the results in terms of both functions and gene expression pathways. GO analysis was able to annotate the results from the GO analysis can interpret genes in terms of cellular component (CC), and molecular function (MF). The analysis was computed using the hypergeometric test and run with the R package clusterProfiler v4.10.0 [16, 17]. Additionally, the visualisation of the enrichment results is implemented using the R package aPEAR [18].

# SMR

SMR can elucidate associations between expression quantitative trait loci (eQTL) and phenotypic outcomes [19]. Specifically, we utilized SMR for our analysis. To assess linkage disequilibrium (LD) and estimate potential co-localization, we performed the HEIDI test using an external reference. Our criteria for defining final significant genes were as follows: (1) SMR FDR: Genes with an SMR FDR P-value < 0.05 were considered significant. (2) Genome-Wide Significance: We required that both eQTL and GWAS results were significantly different at the genome-wide level ( $P < 1 \times 10^{-5}$ ). (3) HEIDI Test Results: Genes were retained if the HEIDI test indicated a result greater than 0.05. These stringent criteria allowed us to identify genes with robust associations related to diabetic retinopathy.

### Selection of cis-eQTL associated with druggable genes

Focusing on the use of cis-expression quantitative trait loci (cis-eQTLs) can better associate with key target genes during drug development. We obtained filtered and significant cis-eQTL data from the eQTLgen consortium [20]. The eQTLgen consortium collected and measured peripheral blood samples from 31,684 projects. These data provide valuable insights into the genetic regulation of gene expression. We also extracted multidimensional genomics data from the NIH-funded PsychENCODE program [21], which focuses on the developing and adult human brain, encompassing both healthy and diseased states, validating the positive genes that extracted from eQTLgen consortium [20].

To identify gene-associated instrumental variables, we selected SNPs within a 100 kb range. These SNPs had an association P-value < 5e-8 with the target gene.  $R^2$  was set as 0.1 to obtain more effective SNPs. Since the original gene database eQTLgen and PsychENCODE was not sufficient to cover existing genes, we selected some genes from the gene database constructed by Vosa U to supplement the MR analysis [20]. The criteria for selecting genes are as follows: (1) All the genes that the P value of TWAS analysis was < 0.05/gene number, the PPH4 > 0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of conditional analysis was < 0.05, the P value of fine-mapping was > 0.9. (2) The genes that the P value of TWAS analysis was < 0.05/gene number, the PPH4>0.8 of colocalization analysis, the P value of permutation test was < 0.05, while their P value of SMR analysis was < 1e–05 with the P value of HEIDI analysis was > 0.05.

In total, 2499 genes were screened from eQTLgen, 1367 genes were included from PsychENCODE and 11 genes (RPS26, SRPK1, WFS1, BABAM1, SUXO, TP53INP1, EIF2S2P3, CCNE2, KRT8P46, LRRC37A15P, SENP2) were selected from Vosa U.

### **Drug-target MR**

Given a sufficient number of high-confidence SNPs, we avoided the use of proxies. To ensure robust instrumental variables, we set the F-statistic threshold at > 10, excluding SNPs with insufficient statistical power [22, 23]. For traits or genes with a single SNP as an instrumental variable, we applied the Wald ratio to assess causal associations. When two SNPs were available, Inverse Variance Weighted (IVW) analysis was employed [24]. For cases with more than two SNPs, we primarily used IVW, supplemented by MR-Egger, Weighted Median, Simple Mode, and Weighted Mode analyses [25–27]. To enhance robustness and reduce potential confounding, we applied the false discovery rate (FDR) correction to p-values. Additionally, we conducted the MR Steiger test for directional validation and investigated the possibility of reverse causality [28]. The MR-Egger intercept test was utilized to detect horizontal pleiotropy. Heterogeneity among the instruments was evaluated using the Cochran's Q test. These MR analyses provided valuable insights into the causal relationships between genetic variants and DR. All analyses adhered to the Strengthening the Reporting of Observational Studies in Epidemiology Using MR (STROBE-MR) guidelines [29]. Furthermore, we complemented our findings with TWAS and SMR analyses, incorporating drug-targeted investigations. All MR analyses were performed using the TwoSampleMR v0.5.11 package in R.

### Integration multiomics evidence of DR-related genes

We combine three main analysis methods to make a comprehensive exploration of DR-related risk genes. Our classifier requires that all candidate genes be related to or mostly related to these three methods. Therefore, we classify candidate genes into three tiers based on specific criteria: Tier 1: The P value of TWAS analysis was < 0.05/ gene number, the PPH4>0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of conditional analysis was < 0.05, the P value of fine-mapping was>0.9, the P value of SMR analysis was<1e-05 with the P value of HEIDI analysis was > 0.05, and the P value of MR verification analysis was significant by the criteria of each gene database. Tier 2: The P value of TWAS analysis was < 0.05/gene number, the PPH4 > 0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of conditional analysis was < 0.05, the P value of fine-mapping was > 0.9, while the P value of SMR analysis was < 1e-05 with the P value of HEIDI analvsis was > 0.05, or the P value of MR verification analysis was significant by the criteria of each gene database. Tier 3: The P value of TWAS analysis was < 0.05/gene number, the PPH4>0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of SMR analysis was < 1e-05 with the P value of HEIDI analysis was > 0.05, and the P value of MR verification analysis was significant by the criteria of each gene database. Tier 4: The other genes that the P value of SMR analysis was < 1e-05 with the P value of HEIDI analysis was >0.05, and the P value of MR verification analysis was significant by the criteria of each gene database. The other genes that the P value of TWAS analysis was < 0.05/gene number, the PPH4 > 0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of conditional analysis was < 0.05, the P value of fine-mapping was > 0.9. The integration of TWAS, colocalization, permutation testing, conditional analysis, fine-mapping, along with SMR and MR, provides a comprehensive and robust framework for identifying potential drug targets for diabetic retinopathy. By leveraging both association and causal inference methods, this approach allows for the precise identification of genes and variants that are not only associated with the disease but also likely to have a direct functional impact on its pathogenesis. The cross-validation between different analytical methods ensures high confidence in the findings, while fine-mapping and MR help to pinpoint the most likely causal variants, significantly increasing the biological relevance of the identified targets. This combined strategy is particularly effective in disentangling confounding factors such as linkage disequilibrium and

pleiotropy, offering a high-resolution view of the genetic mechanisms underlying diabetic retinopathy and thereby facilitating the discovery of viable therapeutic targets.

# **MR-PheWASs**

Functional implications of drug target loci significantly associated with DR (with a significance threshold of P < 0.05) could be explored by employing MR-PheWAS. We selected the top snp of each gene as inputs to explore their association with different trait GWAS summary data on a large scale. The tic analyses were all implemented using the ieugwasr v0.2.2 R package [30]. The biological pathways of the salient genes were queried in the GeneCards database (Supplementary Data 39).

# Metabolomics and immunomics biomarkers data

Metabolome-exposure data from 8229 individuals from the Canadian Longitudinal Study of Aging (CLSA) cohort were used to select 1091 metabolites and 309 metabolite ratios [31]. And as exposure data from 3757 individuals from Sardinia, 731 immune cells (118 absolute cell (AC) counts, 389 median fluorescence intensities (MFI) reflecting surface antigen levels, 32 morphological parameters, and 192 relative cell (RC) counts) related traits were selected [32]. DR-related traits were used as outcomes. We filtered for SNPs associated with 1400 metabolites (P=5E-08) and SNPs associated with 731 immune markers (P=1E-05).  $r^2$  was set as 0.001, and the clump was within 100 kb. These biomarkers were selected to investigate whether there were casual associations between them and DR, which will be of help to us in understand of the deeper mechanisms of the progress of DR (Supplementary Data 30, 32).

### Metabolomics and immunomics MR on DR-related traits

To ensure that there were enough SNPs that could be analysed, we proxied on the 1000 Genomes Project European Reference Sample.  $R^2 = 0.8$  was set as the screening threshold. SNPs in the MHC region were not removed.  $N_{SNP} < =3$  would be removed. TwoSampleMR v0.5.11. was used to conduct a MR analysis from metabolomics and immunomics biomarkers to DR. The method was the same as drug-target MR unless otherwise specified. In addition to Egger intercept test, we used MR-PRESSO as a test for horizontal multiple validity [33]. We exploit the power of multiple genetic variants to consider linkage disequilibrium (LD) between variants. To understand if the associations in terms of genetic variation and outcome were diminished after adjusting for exposure, we carried out leave-one-out analyses in which we progressively removed each SNP, calculated the metaeffects of the remaining SNPs, and observed whether the results changed after removing each SNP. Meanwhile,

we employed heatmaps to represent our aggregated MR results.

### Querying the MGI database

To investigate what happens to our remarkable genes in the mouse knockout model, we interrogated Mouse Genome Informatics (MGI) for significant features revealed [34]. MGI serves as an invaluable international resource for laboratory mice, offering comprehensive genetic, genomic, and biological data. Its standardized nomenclature facilitates the categorization of various mouse strains and their associated features.

# Result

### TWAS revealed significant transcriptomic results with DR

In our investigation, we leveraged GWAS datasets directly or indirectly associated with DR. Employing the FUSION method, along with colocalization and permutation testing, we sought to identify genes significantly linked to DR-related traits (Supplementary Data 1–4). And the summary results could be seen in Supplementary Data 34. Genes exhibiting Z-scores above zero were categorized as up-regulated, while those with Z-scores below zero were considered down-regulated. These regulatory patterns are visually depicted in Fig. 2 (Supplementary Fig. 2–1, 2–2).

Furthermore, we conducted colocalization analysis to ascertain whether the genetic signals associated with these genes and DR traits at a specific locus originate from the same causal polymorphism. Our study identified 23 significant loci associated with DR, 5 features associated with PDR (Supplementary Data 27). All of results passed permutation testing (23/23 in DR, 5/5 in PDR). As it is known, estimated association statistics are well calibrated in the absence of GWAS associations, but may be exaggerated by chance QTL colocalization when GWAS motifs are highly significant and LDs are widespread. Permutation testing rearranges the QTL weights and recalculates the empirical association statistics conditional on the locus GWAS effect. This effectively tests whether identical distributions of QTL effect sizes can produce significant associations by chance, which will help us determine the veracity of the causal associations between these genes and DR-related traits, rather than just resting on the correlation of the huge GWAS data. Additionally, we then performed conditional analysis on the genes corresponding to the key transcription sites (Supplementary Data 5). Next, we used FOCUS fine mapping to pinpoint prospective high-confidence causal genes. 9 features (including RPS26, WFS1, SRPK1, BABAM1, RAB5B and SENP2) in DR and 1 (DCLRE1B) in PDR passed the finemapping test (Table 1, Supplementary Data 6, Supplementary Fig. 2-1, 2-2). This indicates that these genes are



Fig. 2 Circular heat map of TWAS results. A The results of DR. B The results of FDR. The heatmap is divided into five circles with five metrics, namely TWAS-P, COLOC-PPH4, FDR, Conditional Analysis, and FOCUS fine-mapping. The selected molecules have FDR-P values less than 0.05

Phenotype	Tissue	Gene	NEW	TWAS Z score	TWAS P value	FOCUS PIP	Joint P value (conditional analysis)	PP.H4 (colocalization analysis)	Permutation test P value
DR	Pancreas	RPS26	Yes	5.8621	4.57E-09	0.965	4.60E-09	0.99	0.002583
		SRPK1	Yes	4.88043	1.06E-06	0.838	1.10E-06	0.93	0.014019
	Whole blood	RPS26	Yes	5.92052	3.21E-09	0.992	3.20E-09	0.992	0.001371
		WFS1	Yes	5.48151	4.22E-08	1	4.20E-08	0.96	0.00136
	Kidney cortex	RPS26	Yes	5.919	3.24E-09	1	3.20E-09	0.99	0.00911
	sCCA	RAB5B	Yes	5.37171	0.00000078	0.999	7.80E-08	0.908	0.045977
		SENP2	Yes	5.48889	4.04E-08	0.988	4.00E-08	0.916	0.042705
		WFS1	Yes	6.01018	1.85E-09	1	1.90E-09	0.938	0.01872
		BABAM1	Yes	5.50408	3.71E-08	0.991	3.70E-08	1	0.000297
PDR	sCCA	DCLRE1B	Yes	5.15539	2.53E-07	0.896	2.50E-07	0.016	0.01939

Table 1 TWAS significant high confidence genes

High confidence results from TWAS analyses of two DR related phenotypes. TWASs were conducted using cross-tissue expression weights generated from the GTEx v8 release using specific tissues (Pancreas, Whole blood and Kidney cortex) and sparse canonical correlation analysis (sCCA). Significance was defined using a Bonferroni threshold of TWAS P < 0.05/ feature number each tissue after Bonferroni correction (Pancreas of 5734, Whole blood of 7938, Kidney cortex of 1203) and P < 1.32 × 10<sup>-6</sup> (0.05/37,917 cross-tissue sCCA features) in sCCA. Significant TWAS associations were deemed high confidence if they passed a conditional test (joint P value < 0.05) and FOCUS fine-mapping (PIP > 0.5). A gene was defined as novel if it was located greater than 500 kilobases from a lead variant in the source GWAS. Colocalization and permutation analyses were used to further assess the robustness of TWAS findings. TWAS transcriptome-wide association study, PIP posterior inclusion probability, FOCUS Fine-mapping Of CaUsal gene Sets, PP.H4 posterior probability that two traits are associated with a single causal variant, IEAA intrinsic epigenetic age acceleration, GTEx v8 Genotype-Tissue Expression Project version 8, GWAS genome-wide association study

highly DR-associated. The plots of conditional analysis and fine-mapping were showed to observe the superior performance of significant genes (Figs. 3, 4, 5, 6).

### Functional annotations of significant loci

Functional analysis revealed that our TWAS significant genes were mainly associated with intracellular organelle lumen, postsynaptic specialization, cell-substrate junction, endoplasmic reticulum subcompartment, and intracellular vesicle from the perspective of cellular component. And they were also correlated with transcription cis-regulatory region binding, purine ribonucleotide binding, serine hydrolase activity, GTPase regulator activity, peptidase inhibitor activity, channel activity and, carboxylic acid transmembrane transporter activity from molecular function (Fig. 7, Supplementary Data 7).



Fig. 3 Conditional analysis plots for key positive genes for specific tissues. A RPS26 in pancreas. B RPS26 in whole blood. C SRPK1 in pancreas. D WFS1 in whole blood. E RPS26 in kidney cortex. These are the genes that characterize DR. Conditioned/marginal features were significantly associated with DR only in the unadjusted model. After correction, independently/jointly significant features remained associated with the phenotype at a nominal significance level (p < 0.05)

# High-confidence SMR multi-tissue findings for DR

Within the range of 49 tissues analysed in the transcriptome, we performed a large summary data-based Mendelian randomization (SMR) analysis comprehensively covering most of the possibilities, looking for 92 key DRregulated genes that were significantly positive in 49 tissues a total of 446 times (Table 2, Supplementary Data 8, 9). The screening thresholds were SMR P value < 1E-05along with FDR P value < 0.05, HEIDI test P value > 0.05. Sixty of these genes are protein-codable. Of the genes, the most significantly related in terms of association are the HLA-DQ family genes, with HLA-DQB1, B2, HLA-DQA2 showing the top three significance, the latter two being the most significant in the blood, and the topmost being the most specifically and significantly expressed in sun-exposed skin. From the perspective of gene overlap, the most frequent occurrences of our genes were the HLA-DQ family. What excited us is that RPS26 and WFS1 are two novel co-significant genes revealed in multi DR weights corresponding with the result of TWAS analyses (Tables 1, 2). We identified that the genetic loci of RPS26 and WFS1 are in close proximity, suggesting a potential combined influence on the development of DR (Fig. 8). Furthermore, SUXO exhibited significantly positive SMR\_P values (Fig. 9), appearing across 14 different tissue types. These three genes constitute the most pivotal elements identified in the SMR analysis, highlighting their critical relevance to the findings.

GO enrichment analysis made a co-ordinated statistical analysis of the functions of the significant genes. The results showed that numerous genes collectively pointed to eight immune-related pathways, which suggested a close relationship between the progression of DR and immune molecules, especially the assembly of the MHC class II protein complexes among them, which may be a key link in the process of DR development (Supplementary Data 10, 11).

# High-confidence SMR multi-tissue findings for PDR

A total of 327 positive results were seen in 49 tissues containing 55 genes (Table 3, Supplementary Data 12, 13). The screening conditions were as in the previous section. A total of 37 protein-codable genes were included. In terms of significance results, the top five were all HLA-DQ family: HLA-DQB1, B2, B6, A1 and A2.In terms of frequency of occurrence, a total of 13 genes appeared with a frequency of greater than or equal to five times, of which, in addition to the top ranked HLA-DQB2, which was the most prominently expressed in the blood, and the results of other HLA-DQ family genes, there was also the result of SKIV2L expressed in the Artery\_Aorta with a total of 23 occurrences. Also, the gene TNXA is a locus of interest with a total frequency of 15 occurrences.



Fig. 4 Fine-mapping diagrams of key genes. A RPS26 (ENSG00000197728.9, 12:56041351-56044697) in DR from pancreas. B RPS26 (ENSG00000197728.9, 12:56041351-56044697) in DR from whole blood. C RPS26 (ENSG00000197728.9, 12:56041351-56044697) in DR from kidney cortex. RPS26 is a hub gene in regulation of DR, which was significant in multiple analysis results. D WFS1 (ENSG00000109501.13, 4:6269849-6303265) in DR from whole blood. E SRPK1 (ENSG0000096063.15, 6:35832966-35921342) in DR from pancreas. All of these genes pass the fine-mapping exam. It means that they have a casual effect of the regulation of DR. The direction of their effect can be analysis by the TWAS-Z score



**Fig. 5** Conditional analysis plot of significant key genes in the tissue-wide database sCCA. **A** WFS1. **B** RAB5B. **C** SENP2. **D** DCLRE1B. **E** BABAM1. Among them, WFS1, RAB5B, SENP2 and DCLRE1B are characteristic genes of DR, and DCLRE1B is the important characteristic gene of PDR, distinguishing it from DR. Conditioned/marginal features were significantly associated with DR only in the unadjusted model. After correction, independently/jointly significant features remained associated with the phenotype at a nominal significance level (p < 0.05)

We explored the functional annotations of related genes on the GO database and found nine immunerelated functional annotations including MUC class II protein complex, myeloid dendritic cells and antigenantibody binding (Supplementary Data 14, 15).

# **Drug-target MR**

We extracted and screened cis-eQTL data from each of the two databases, eQTLgen, Psychencode, and ended up with 2499 and 1367 corresponding genes (Supplementary Data 16, 17). We did a large-scale drug target MR analysis of them with outcomes of DR and PDR respectively. The p-value thresholds for MR analyses were determined using a Bonferroni correction approach, calculated as 0.05 divided by the total number of genes analyzed, resulting in thresholds of MR p-value < 0.00002001 for eQTLgen (with 2499 genes), p-value < 0.00003658 for Psychencode (with 1367 genes) and p-value < 0.00455 for Vosa U (with 11 genes). FDR was also used to observe the corrected change in P values as a primary screening index. Q test P value of heterogeneity>0.05 and Egger intercept test P value > 0.05 were applied as sensitive test (Table 4). The Steiger filter is used to check the directionality of MR results in eQTLgen (Supplementary Data 20, 25).

We obtained a range of 14 significant genes in the eQTLgen with DR as well as 13 with PDR (Supplementary Data 19, 24). Shared genes between DR and PDR

amounted to five (ITGB7, TEK, ITPR3, C4A, PRKD2, ERBB3) in eQTLgen consortium in primary screening (P<0.00002001 for IVW) (Fig. 10). After colocalization analysis, a total of 11 high confidence features were selected from eQTLgen (RPS26 of DR, EIF2S2P3 of DR, KRT8P46 of DR, LRRC37A15P of DR, TP53INP1 of DR, CDH2 of DR, CTLA4 of DR, ITGB7 of DR, ERBB3 of DR, KSR1 of PDR, ITGB7 of PDR) (Table 4, Supplementary Data 18, 37). Gene ERBB3 of PDR in eQTLgen was also closed to positive with a COLOC.PP.H4 of 0.949, which might be a hub gene in the regulation of DR and PDR. When performing a double test using Psychencode for DR, we found that two genes (SRPK1, ERBB3) were positive for both eQTLgen and Psychencode in primary screening (Supplementary Data 18). In the Psychencode's test, both the ERBB3 and HLA-B genes showed positive results in both DR and PDR (Supplementary Data 21, 26). All of these results suggested that ERBB3 and SRPK1 might be candidate effector genes for DR and PDR susceptibility. The results of Egger intercept test and Q heterogeneity test could be seen in Supplementary Data 22, 23, 27, 28.

### Integrated multi-omics evidence of DR-related genes

We integrated and analyzed the results of the three main methods. All significant genes are divided into four tiers (Table 5, Supplement Data 36). We found that the RPS26 gene in Whole blood showed



Fig. 6 Fine-mapping diagrams of key genes. A WFS1 (ENSG0000109501.13, 4:6269849-6303265) in DR from sCCA. It is a new discovery that marks the mechanism of DR occurrence, which also corresponding with results in TWAS. B RAB5B (ENSG00000111540.15, 12:55973913-55996683) in DR from sCCA. C BABAM1 (ENSG00000105393.15, 19:17267376-17281249) in DR from sCCA. D SENP2 (ENSG00000163904.12, 3:185582496-185633551) in DR from sCCA. E DCLRE1B (ENSG00000118655.4, 1:113904619-113914086) in PDR from sCCA. It is the only essential gene found to distinguishes DR from PDR. All of these genes pass the fine-mapping exam. It means that they have a casual effect of the regulation of DR. The direction of their effect can be analysis by the TWAS-Z score

significant positive effects on DR development in all three methods (TWAS.Z=5.92052, SMR.Beta=0.071, MR.Beta=0.180). These suggest that there is a causal effect between RPS26 and the development of DR. WFS1 and SRPK1 (Tier 2) also show a causal relationship with the development of DR (TWAS.Z.<sub>WFS1</sub>=5.4815, SMR.Beta.<sub>WFS1</sub>=0.157814, TWAS.Z.<sub>SRPK1</sub>=4.88043,

MR.Beta.<sub>SRPK1</sub>=0.124). In the tier 3: TP53INP1, KRT8P46 and LRRC37A15P in Whole blood showed a positive effect on the development of DR in TWAS, SMR and MR. While EIF2S2P3 in Whole blood shows a negative effect on the regulation of DR in TWAS, SMR and MR. A controversial point is that the TWAS results of the CCNE2 gene show positive regulation



Fig. 7 Functional analysis of significant causal genes for DR-related traits. A Cellular component of key genes. B Molecular function of key genes

Gene	No.of Tissue	Result in the	Category				
		β	SE	P value	Tissue		
HLA-DQB2	40	0.425821	0.0180877	1.5158E-122	Whole_Blood	PR	
HLA-DQB1	37	- 0.38162	0.0154173	2.904E-135	Skin_Sun_Exposed_Lower_leg	PR	
RPS26	36	0.0701062	0.0123345	1.31778E-08	Nerve_Tibial	PR	
HLA-DQA2	35	0.403143	0.0219115	1.34585E-75	Whole_Blood	PR	
HLA-DRB6	32	0.417842	0.0229237	3.12088E-74	Muscle_Skeletal	PS	
HLA-DQA1	30	- 0.601957	0.0393946	1.03619E-52	Skin_Sun_Exposed_Lower_leg	PR	
HLA-DRB1	25	- 0.882803	0.0619456	4.40165E-46	Lung	PR	
SUOX	14	- 0.258211	0.0440463	4.5658E-09	eQTLGen	PR	
HLA-DQB1-AS1	13	- 0.508678	0.0512114	2.994E-23	Testis	PR	
SKIV2L	12	- 0.345387	0.0389783	7.93073E-19	Pancreas	PR	
RPL32P1	10	- 0.212719	0.0254583	6.51015E-17	Skin_Sun_Exposed_Lower_leg	PS	
LEMD2	10	- 0.506115	0.0915501	3.2339E-08	Heart_Left_Ventricle	PR	
HCG17	9	- 0.212533	0.0351113	1.42069E-09	Adrenal_Gland	RNA Gene (IncRNA)	
WFS1	8	0.157814	0.0278138	1.39537E-08	Skin_Not_Sun_Exposed_Suprapubic	PR	
TNXA	7	- 0.344274	0.0363142	2.53255E-21	Adipose_Subcutaneous	PS	
UQCC2	7	- 0.525417	0.0777439	1.39613E-11	Skin_Sun_Exposed_Lower_leg	PR	
NELFE	6	0.957561	0.152764	3.65149E-10	Skin_Not_Sun_Exposed_Suprapubic	PR	
HLA-DRB5	5	- 0.41296	0.0382841	3.97741E-27	Ovary	PR	
HLA-J	5	0.14534	0.0253651	1.00486E-08	Artery_Aorta	PS	

Table 2	High-cor	nfidence S	SMR mu	lti-tissue	gene	results	for DR

SMR results for DR. Summary data-based Mendelian randomization (SMR) analysis were conducted, looking for 92 key DR-regulated genes that were significantly positive in 49 tissues a total of 446 times. The screening thresholds were SMR P value < 1E–05 along with FDR P value < 0.05, HEIDI test P value > 0.05. Genes with high significance in 5 or more tissues are shown. We defined protein-coding according to HGNC, Ensembl, or NCBI Gene by GeneCards. RNA gene according to HGNC, Ensembl, or NCBI Gene or genes that are mined from RNAcentral and its external sources are defined as ncRNA genes. Pseudogene were defined according to HGNC, Ensembl, or NCBI Gene. All of the Category were obtained in https://www.genecards.org/\_PR Protein Coding; PS Pseudogene

(TWAS.Z=6.17164), and the SMR and MR results show a negative causal relationship (SMR.Beta=--0.176809, MR.Beta=--0.2664). In the tier 4: EHMT2, C4A, ERBB3 and SUOX show a negative effect on the development of DR. SENP2 and RAB5B show a positive effect on the regulation of DR.



Fig. 8 SMR LOCUS Plot of RPS26 and SUXO. These two genes (RSP26 and SUXO) inherited together with their locus nearby. They have a significant correlation with DR, which validates the casual relation with DR in TWAS analysis again

# **MR-PheWAS**

To further explore the potential impact of our significant genes, we did MR-PheWAS analysis using top SNPs for each gene. p-values less than 1E-05 were defined as positive. The results showed that three out of 12 features duplicated the top SNP: rs223490 (LRRC37A15P, KRT8P46 and RP11-10L12.2). This snp shows its potential role includes the effects of elevating  $\beta$ -mannosidase, lowering cystatin C levels, lowering serum creatinine, and elevating glomerular filtration rate levels. The causal effect of EIF2S2P3 and WFS1 in DR shown by PheWAS was to induce the onset of type 2 diabetes. And RPS26 and SUOX in DR were to induce rheumatoid arthritis and reduce asthma. And these two genes along with EIF2S2P3 and SRPK1 in DR can also reduce the level of eosinophil count. And rs223490 (LRRC37A15P, KRT8P46 and RP11-10L12.2), RPS26 and SUOX elevate the level of lymphocyte count in DR. The total effects of these genes point to predisposing factors of DR. Detailed results could be viewed in Supplementary Data 29. It also points out that DR might associated with a range of different diseases through the effects of different genes, which need further studies in the future.

# Integrative metabolomic and immunomic insights into the progression of DR

Given the prevalence of positive associations in blood tissue and immune molecules in SMR analysis, we propose employing immunohistology and metabolomics to further investigate the identified genes and uncover potential biomarkers. Positive results in metabolomics and immunomics were defined using the following: P value of IVW <0.05, and positive FDR test as well as Q teat P value of heterogeneity >0.05, Egger\_intercept test P value >0.05, and also a p-value of >0.05 for MR-PRESSO. At the same time, we used the MR Egger to test for horizontal multiple validity and its P-value should also be greater than 0.05. After our multiple sensitivity test screening, we screened the positive results with high confidence (Supplementary Data 31, 33).

We obtained 38 positive results from 1,400 metabolites, the first five known metabolite positive results were: 1-stearoyl-GPG (18:0) levels ( $\beta = -0.255$ ), 1-palmitoyl-GPE (16:0) levels ( $\beta = -0.224$ ), Caffeine levels ( $\beta = 0.434$ ), 1-stearoyl-2-oleoyl-GPE (18:0/18:1) levels ( $\beta = -0.149$ ), Hexanoylglutamine levels ( $\beta = 0.169$ ) (Table 6). Most of these identified molecules are rarely studied molecules,



Fig. 9 SMR LOCUS Plot of WFS1. WFS1 is an essential gene in the regulation of DR. It has a significant correlation with DR, which validates the casual relation with DR in TWAS analysis again

Gene	No.of Tissue	Result in the	Category			
		β	SE	P value	Tissue	
HLA-DQB2	41	0.355233	0.0178844	8.55356E-88	Whool Blood	PR
HLA-DQB1	37	- 0.318443	0.0154726	4.05585E-94	Skin_Sun_Exposed_Lower_leg	PR
HLA-DRB6	32	0.404109	0.0240224	1.6803E-63	Muscle Skeletal	PS
HLA-DQA1	30	- 0.58236	0.0404634	5.78595E-47	Skin_Sun_Exposed_Lower_leg	PR
SKIV2L	23	- 0.205019	0.0302312	1.1876E-11	Artery_Aorta	PR
HLA-DRB1	22	- 0.907715	0.075105	1.25313E-33	Thyroid	PR
HLA-DQB1-AS1	19	- 0.378579	0.0358677	4.8227E-26	Putuitary	RNA Gene (IncRNA)
TNXA	15	- 0.327415	0.05003	5.97452E-11	Skin_Sun_Exposed_Lower_leg	PS
RPL32P1	11	- 0.132668	0.0249025	9.95779E-08	Skin_Sun_Exposed_Lower_leg	PS
HLA-DRB5	8	- 0.279709	0.0261476	1.04771E-26	Liver	PR
NELFE	7	0.973061	0.18913	2.67607E-07	Muscle Skeletal	PR
C4A	6	- 0.296959	0.0338264	1.65056E-18	Esophagus_Mucosa	PR
UQCC2	5	- 0.387118	0.0719026	7.28782E-08	Skin_Sun_Exposed_Lower_leg	PR

Table 3 High-confidence SMR multi-tissue gene results for PDR

SMR results for PDR. SMR analysis were conducted, looking for 55 key PDR-regulated genes that were significantly positive in 49 tissues a total of 327 times. The screening thresholds were SMR P value < 1E–05 along with FDR P value < 0.05, HEIDI test P value > 0.05. Genes with high significance in 5 or more tissues are shown. We defined protein-coding according to HGNC, Ensembl, or NCBI Gene by GeneCards. RNA gene according to HGNC, Ensembl, or NCBI Gene or genes that are mined from RNAcentral and its external sources are defined as ncRNA genes. Pseudogene were defined according to HGNC, Ensembl, or NCBI Gene. All of the Category were obtained in https://www.genecards.org/. PR Protein Coding; PS Pseudogene

Gene	Phenotype	Source	method	snp	Beta	SE	pval	FDR-adjusted P value	COLOC.PP.H4
RPS26	DR	Vosa U	MR Egger	14	0.148	0.068	0.048	2.06E-07	0.999
			Weighted median		0.181	0.031	5.98E-09		
			IVW		0.180	0.034	9.38E-08		
EIF2S2P3	DR	Vosa U	MR Egger	10	- 0.428	0.119	0.007	1.49E-09	0.999
			Weighted median		- 0.287	0.052	2.76E-08		
			IVW		- 0.266	0.042	2.70E-10		
KRT8P46	DR	Vosa U	MR Egger	16	0.115	0.067	0.108	4.65E-08	0.997
			Weighted median		0.159	0.037	9.44E-06		
			IVW		0.151	0.027	1.69E-08		
LRRC37A15P	DR	Vosa U	MR Egger	23	0.056	0.056	0.332	2.21E-09	0.996
			Weighted median		0.151	0.031	1.03E-06		
			IVW		0.144	0.023	6.02E-10		
TP53INP1	DR	Vosa U	MR Egger	27	0.116	0.075	0.134	1.45E-16	0.999
			Weighted median		0.192	0.038	3.02E-07		
			IVW		0.225	0.026	1.32E-17		
CDH2	DR	eQTLgen	MR Egger	28	0.059	0.058	0.314	0.003	0.833
			Weighted median		0.151	0.038	7.17E-05		
			IVW		0.127	0.030	2.00E-05		
CTLA4	DR	eQTLgen	MR Egger	6	- 0.342	0.282	0.292	0.002	0.974
			Weighted median		- 0.474	0.099	1.80E-06		
			IVW		- 0.340	0.075	6.40E-06		
ITGB7	DR	eQTLgen	MR Egger	10	0.104	0.043	0.040	0.003	0.738
		-	Weighted median		0.110	0.034	0.001		
			IVW		0.109	0.026	2.13E-05		
ERBB3	DR	eQTLgen	MR Egger	4	- 0.838	0.484	0.226	5.37E-07	0.949
		-	Weighted median		- 0.478	0.087	3.53E-08		
			IVW		- 0.479	0.078	7.58E-10		
KSR1	PDR	eQTLgen	MR Egger	19	- 0.138	0.054	0.021	0.0004	0.611
		-	Weighted median		- 0.145	0.037	9.50E-05		
			IVW		- 0.131	0.026	7.79E-07		
ITGB7	PDR	eQTLgen	MR Egger	10	0.104	0.045	0.050	0.002	0.852
			Weighted median		0.145	0.035	2.77E-05		
			IVW		0.120	0.027	8.90E-06		

Table 4 High confidence genes of drug-target MR

Drug-target MR results of DR. Data was sourced from eQTLgen and Vosa U studies. IVW Inverse variance weighted

perhaps to be added to future studies (Fig. 11). More detailed can be seen in Supplementary Data 31.

The MR analysis of 731 immunomarkers with DR and obtained 26 positive results according to the MR result positivity test described above. Among them, several clusters of key interest were demonstrated by molecular clusters expression on different immune cells predominantly: expression of RAFF-R on different immune cells  $(\beta < 0)$ , prominent expression of CD3  $(\beta < 0)$ , and expression of CD80 molecules on dendritic cells and monocytes  $(\beta > 0)$ . The top five prominent immune markers were Plasmacytoid Dendritic Cell Absolute Count, Activated CD4 regulatory T cell Absolute Count, Central Memory CD8+T cell %T cell, Naive CD8+T cell %CD8+T cell and CD4+T cell Absolute Count. But disappointingly after FDR correction most of the molecules were shown to be insignificant. More detailed can be seen in Supplementary Data 33.

# Mouse knock-out models for novel genes identified by TWAS or drug-target MR

We interrogated the Mouse Genomics (MGI) resource for relevant trait expression in knockout mice. In this way, we reached to demonstrate direct post-experimental validation of correlation between key genes and DR. Among the 12 key genes searched, we found WFS1



Fig. 10 Results of drug-target MR in eQTLgen consortium. A Volcano plot of DR results from Mendelian randomization of drug targeting. B Volcano plot of PDR results from Mendelian randomization of drug targeting. C Plots of colocalization results of CDH2 in DR for Mendelian randomization of drug targets. D Plots of colocalization results of CTLA4 in DR for Mendelian randomization of drug targets. E Plots of colocalization results of ITGB7 in DR for Mendelian randomization of drug targets. F Plots of colocalization results of ERBB3 in DR for Mendelian randomization of drug targets. G Plots of colocalization results of KSR1 in PDR for Mendelian randomization of drug targets. H Plots of colocalization results of ITGB7 in PDR for Mendelian randomization of drug targets.

Grading	Genes	Tissue	TWAS.P	TWAS.Z	PP.H4	PIP	SMR.P	SMR.Beta	MR_IVW.P	MR.Beta
Tier 1	RPS26	Whole Blood	3.21E-09	5.92052	0.992	0.992	1.318E-08	0.0701062	9.38E-08	0.180357536
Tier 2	WFS1	Whole Blood	4.22E-08	5.48151	0.96	1	1.395E-08	0.157814	0.02107	0.150781049
	SRPK1	Pancreas	1.06E-06	4.88043	0.93	0.838	NA	NA	2.67E-05	0.124082046
Tier 3	TP53INP1	Whole Blood	1.06E-07	5.3157	0.982	NA	5.54287E-08	0.311773	1.32E-17	0.225211828
	EIF2S2P3	Whole Blood	4.23E-07	- 5.06	0.791	NA	8.23845E-06	- 0.176809	2.70E-10	- 0.266404116
	CCNE2	sCCA	6.76E-10	6.17164	1.00	NA	2.58816E-07	- 1.18355	0.0879	- 0.512961065
	KRT8P46	Whole Blood	5.01E-06	4.56437	0.965	NA	6.15695E-06	0.115008	1.69E-08	0.151250504
	LRRC37A15P	Whole Blood	5.01E-06	4.56437	0.965	NA	5.84132E-06	0.1018	6.02E-10	0.144487695
Tier 4	EHMT2	eQTLGen	NA	NA	NA	NA	8.53393E-11	- 2.57416	2.45E-103	- 2.654362913
	C4A	Cells_Cul- tured_fibro- blasts	NA	NA	NA	NA	3.04948E-19	0.547908	5.72E-14	- 0.668123099
	ERBB3	eQTLGen	NA	NA	NA	NA	1.37764E-08	- 0.507111	7.58E-10	- 0.478940464
	SUOX	Pancreas	1.34E-07	- 5.2729	1.00	NA	4.5658E-09	- 0.258211	0.001066	- 0.102746816
	BABAM1	sCCA	3.71E-08	5.50408	1.00	0.991	NA	NA	3.53E-04	- 0.156962244
	SENP2	sCCA	4.04E-08	5.48889	0.916	0.988	NA	NA	0.60281	- 0.077820202
	RAB5B	sCCA	7.80E-08	5.37171	0.908	0.999	NA	NA	NA	NA

 Table 5
 The grading table of article results

Tier 1: The P value of TWAS analysis was < 0.05/gene number, the PPH4 > 0.8 of colocalization analysis, the P value of permutation test was < 0.05, the P value of fine-mapping was > 0.9, the P value of SMR analysis was < 1e–05 with the P value of HEIDI analysis was < 0.05, the P value of fine-mapping was > 0.9, the P value of SMR analysis was < 1e–05 with the P value of HEIDI analysis was < 0.05, the P value of FINE analysis was < 0.05, the P value of FINE analysis was < 0.05, the P value of SMR analysis was < 0.05, the P value of FINE analysis was < 0.05, the P value of SMR analysis was < 0.05, the P value of FINE analysis was < 0.05, the P value of fine-mapping was > 0.9, while the P value of SMR analysis was < 0.05, the P value of FINE analysis was < 0.05, the P value of SMR analysis was < 0.05, the P value of SMR analysis was < 1e–05 with the P value of HEIDI analysis was > 0.05, or the P value of MR verification analysis was significant by the criteria of each gene database. Tier 3: The P value of TWAS analysis was < 0.05/gene number, the PPH4 > 0.8 of colocalization analysis, the P value of permutation test was < 0.05, gene number, the PPH4 > 0.8 of colocalization analysis, was significant by the criteria of each gene database. Tier 3: The P value of TWAS analysis was < 1e–05 with the P value of HEIDI analysis was > 0.05, and the P value of MR verification analysis was < 1e–05 with the P value of HEIDI analysis was > 0.05, and the P value of MR verification analysis was < 1e–05 with the P value of SMR analysis was < 1e–05 with the P value of SMR analysis was < 1e–05 with the P value of MR analysis was > 0.05, and the P value of MR verification analysis was < 0.05, and the P value of MR verification analysis was < 1e–05 with the P value of SMR analysis was < 0.05, and the P value of MR verification analysis was < 0.05, and the P value of MR verification analysis was < 0.05, and the P value of MR verification analysis was < 0.05, and the P value of MR verification analysis was < 0.05, and the P value of MR

directly associated with DR. The annotated information we found to be highly correlated with significant gene presentations included abnormal insulin levels, abnormal cellular function, and abnormal reproductive system function. (Supplementary Data 35).

# Discussion

Our study utilized extensive genomic data of DR and PDR, employing a multi-omics approach structure to elucidate the relationship between various drug-target features and DR-related traits (Fig. 1). In addition to identifying seven positively associated genes (RPS26, SRPK1, WFS1, RAB5B, SENP2, BABAM1) linked to DR, we specifically highlighted DCLRE1B as a gene unique to PDR. These genes are associated with conditions such as Wolfram Syndrome [35], anemia [36, 37], and non-alcoholic steatohepatopathy [38]. Notably, RPS26, WFS1 and SRPK1 exhibited significant causal effects in the validation processes of SMR and drug-targeted MR (Tables 2, 4). Beyond their association with DR, we focused more on their pleiotropic effects,

delving into the multi-trait impacts of these genes. The loci and proteins identified in this study represent promising pharmacological targets, thereby enhancing the evidence hierarchy in the drug discovery process and laying a foundational basis for subsequent drug research and its translational applications.

# Key genes RPS26, WFS1, and SRPK1: exploring their roles and mechanisms in diabetic retinopathy

Among the genes co-expressed in our SMR and TWAS, a remarkably important positive gene appeared: the RPS26. It appeared positive in TWAS four times in different tissues (Both 3 specific tissues and sCCA showed positivity), as well as manifested in SMR with positive results in 36 tissues. The most prominent tissues were whole blood and nerve tibial respectively. Remarkably, RPS26 mediates the ribosomal stress response, and in the face of stress, the cysteines in RPS26 and Rpl10 are readily oxidized and then released from the ribosome via their chaperones Tsr2 and Sqt1, followed by targeted repair of the damaged ribosome with newly made proteins [39,

Phenotype	Exposure	IVW_Beta	IVW_SE	IVW_P	IVW_P_FDR	MR Egger_pval	Weighted median_ pval	Weighted mode_pval
Diabetic retinopathy	1-stearoyl-GPG (18:0) levels	- 0.255	0.067	< 0.001	< 0.001	0.957	0.000145309	0.132
	1-palmitoyl-GPE (16:0) levels	- 0.224	0.064	< 0.001	< 0.001	0.526	0.000411367	0.024
	X-16087 levels	- 0.204	0.048	< 0.001	< 0.001	0.096	3.71659E-06	0.014
	Caffeine levels	0.434	0.135	0.001	0.028538	0.499	0.005268488	0.22
	1-stearoyl-2-oleoyl-GPE (18:0/18:1) levels	- 0.149	0.047	0.001	0.028538	0.124	0.000604548	0.03
	Hexanoylglutamine levels	0.169	0.049	0.001	0.028538	0.213	1.66621E-06	0.039
	Deoxycholic acid glucuronide levels	- 0.134	0.039	0.001	0.028538	0.277	0.002586349	0.101
	Tetradecanedioate (C14-DC) levels	- 0.097	0.03	0.001	0.028538	0.375	0.003026664	0.092
	X-25371 levels	- 0.243	0.075	0.001	0.028538	0.214	7.7527E-07	0.021
	1-linoleoyl-GPG (18:2) levels	- 0.12	0.035	0.001	0.028538	0.754	0.00098631	0.134
	Octadecenedioylcarni- tine (C18:1-DC) levels	- 0.089	0.028	0.001	0.028538	0.234	0.002113406	0.088
	1-stearoyl-GPE (18:0) levels	- 0.194	0.062	0.002	0.043647	0.575	0.001295219	0.021
	Hexadecenedioate (C16:1-DC) levels	- 0.082	0.026	0.002	0.043647	0.754	0.000768446	0.142
	Deoxycholic acid 12-sulfate levels	- 0.119	0.039	0.002	0.043647	0.731	0.0052984	0.111
	Hexadecanedioate (C16-DC) levels	- 0.091	0.029	0.002	0.043647	0.253	0.01100617	0.102

Table 6	MR revea	ls causal re	elationship	between	metabo	plomics and	DR
---------	----------	--------------	-------------	---------	--------	-------------	----

Results of MR analyses for DR and metabolomics. Genes with both IVW\_P values and FDR\_P values less than 0.05 were included as positive high-confidence results. MR Mendelian randomization, FDR false discovery rate, IVW inverse variance weighted

40]. We must emphasize that oxidative stress is one of the key mechanisms of diabetes and diabetic retinopathy. Oxidative stress of RPS26 leads to ribosomal damage and is likely to play a special role in the process that leads to the occurrence of DR. RPS26 mediates the formation of anemia, which may be one of the reasons why it is most prominent in whole blood. As we know, diabetic patients often develop anemia, this may be one of the important co-regulated genes [36, 37]. However, it has been shown that RPS26 is more strongly expressed in multiple tissues, it means that it is more likely to affect a number of different phenotypes with some pleiotropy [41]. This suggests that the mechanism by which RPS26 works may be pervasive in multiple tissues throughout the body. Pathway analysis found that it was strongly associated with both Pre-mRNA and rRNA (Supplement Data 39). Therefore, it is very likely that the mechanism of gene regulation will be implemented. Simultaneously, it is important to note the close genetic relationship between SUXO and RPS26, as evidenced by numerous overlapping SNP variants observed in the regional location map. From our pathway analysis, RPS26 and SUXO share two biological pathways: Metabolism and Nervous system development. It is reasonable to assume that neurometabolites are likely to be mediators of the interaction or co-action of the two. Further research is required to explore their regulatory interactions in relation to DR.

Except for RPS26, SMR analyses also highlighted the conclusiveness of positivity for WFS1. WFS1, a high confidence DR gene, associated with Wolfram Syndrome, which is characterized by juvenile-onset diabetes and diabetes insipidus [42]. It binds directly to vesicular cargo proteins, including insulinogen, through the C-terminal segment of its endoplasmic reticulum lumen, and its deficiency leads to an abnormal accumulation of insulinogen in the endoplasmic reticulum, impeding insulinogen processing and insulin secretion [43]. Transcriptomic studies using the WFS1 mouse model have revealed the molecular pathways affected after WFS1 gene knockout, including the G protein signaling pathway, the ER stress pathway, and the proteasome/lysosomal pathway. In pancreatic cells of WFS1 mice, insulin secretion is reduced, which is associated with a decrease in TRPM5 gene expression



Fig. 11 Circular heatmap of positive Mendelian randomisation results on initial screening for metabolomics and immunomics. Results with IVW\_P values < 0.05 were included. We consider that these results suggest that there may be a causal association of these numerous molecules with DR, especially a certain few molecules that occur in clusters in multicellularity

[44]. WFS1 was highly correlated with type 2 diabetes-related traits, which is the most closely related. To further understand the biology behind it. Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs) and CAMKK2 pathway might be its pathways to DR (Supplement Data 34, 39).

Our drug-targeted MR analysis revealed numerous key genes. SRPK1 (serine-rich protein kinase-1) promotes angiogenesis by phosphorylating serine-rich splicing factor-1 (SRSF1), a regulator of vascular endothelial growth factor splicing [45]. Inhibition of SRPK1 contributes to selective down-regulation of pro-angiogenic VEGF isoforms [46, 47]. Therefore, SRPK1 inhibitors can be delivered as ophthalmic drops to diminish retinal permeability and edema in the DR [48, 49]. In addition to this, SRPK1 can selectively splicing exacerbate non-alcoholic steatohepatopathy through SRSF6-related RNA [38]. It focuses on the pro-metastatic effect on epithelial hepatocellular carcinoma cells and the pro-productive effect on mesenchymal stroma [50]. As for SRPK1, we could discover that it corelated with mean corpuscular hemoglobin and eosinophil cell count. We discovered that it contributes to RNA Polymerase I Promoter Opening, Packaging of Telomere Ends, Processing of Capped Intron-Containing Pre-mRNA and VEGFA-VEGFR2 signaling (Supplement Data 34, 39). Our analysis highlights genes like RPS26, WFS1 and SRPK1 as essential targets for future therapeutic strategies aimed at diabetic retinopathy. By developing treatments that specifically target these genes, we can address the underlying causes of the disease more effectively. This approach not only aligns with the pressing need for improved treatment options but also enhances the potential for personalized medicine in managing DR.

# Other mechanistic insights into DR: central roles of SUMO, DCLRE1B, and PRKD2

TWASs analysis revealed 9 meaningful features in different tissues that are associated with DR and 1 with PDR (Table 1). BABAM1 was identified as a key gene associated with breast cancer [51]. It regulates damagedependent BRCA1 localization by acting early in the DNA damage response [52]. This may be a potential DRrelated mechanism. RAB5B is a regulator of exosome secretion, which may contribute to the development of DR by promoting exosome secretion [35–54]. SUMOspecific protease 2 (SENP2) is a de-SUMO-enzymes. Our results suggest that its elevated expression may promote DR leading to endothelial dysfunction and atherosclerosis [55]. Most of the other initial screen-positive genes were excluded during fine-mapping and conditional analyses, but they still need to be studied further. Aging regulator DCLRE1B may serve as a candidate effector gene for T1D susceptibility [56]. It uses its activity to help remove common oxidative damage in telomeres and has a good leading telomere excision effect [57, 58]. As a key gene that distinguishes DR from PDR, it may induce an exacerbation of disease severity through the mechanism of telomere resection.

In addition to these, there is a special gene to focus on. PRKD2 (protein kinase D) is a biomolecule with multidirectional functions. It showed multiple positives or nearpositive results in our multiplex analysis, which suggested that it might also be a hub gene. A nonsense mutation in the gene for PRKD2 was found in rhesus monkeys with very high insulinemia. In PRKD2-KO mice, the deletion/ downregulation of PRKD2 was found to mediate hyperinsulinemia, leading to insulin resistance (IR) and metabolic disorders [59]. And IR is the primary disease that mediates the formation of DR [60]. This is consistent with our TWAS and DrugMR results: PRKD2 reduces the risk of DR ( $\beta$  < 0), and when PRKD2 is down-regulated, the probability of DR onset is greatly increased. In addition to this, PRKD2 has a correlation for tumor growth and angiogenesis, which coincides with the possible mechanism of DR pathogenesis [61]. And it has been shown that PRKD2 can also control neutrophil differentiation by participating in the HAX1-dependent control of mitochondrial protein homeostasis dysregulation, which fully illustrates the possibility of immune-mediated mechanisms involved, and fits with our immune exploration [62].

ERBB3 acts as a regulator of cytoskeletal dynamics in microvascular endothelial cells, affecting vascular endothelial permeability and tight junction levels [63], suggesting that MSC-CM improves impaired vascular endothelial permeability in diabetic patients by regulating ERBB3. This indicated that the results that ERBB3 has a protective effect on the occurrence of DR and PDR is credible and has a research basis, which is promising to continue to explore it in depth.

In the SMR analysis, significant positive gene expression for chromosome 6 can first be observed. The cooccurrence of HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2 characterizes the importance of the immune milieu for the development of DR. HLA molecules have been shown to be positive for diabetes-inducing properties in European and African populations, and our results reaffirm the reliability of this result [64, 65]. In terms of tissue specificity, when we excluded the results of chromosome 6, the SMR results of PDR were excluded, leaving only partial DR results. Among the most frequent tissues were blood (eQTLgen) and Skin\_Sun\_Exposed\_ Lower\_leg. The top five genes with the highest number of positives were RPS26, SUOX, SKIV2L, RPL32P1 and LEMD2. Together, these genes were shown to regulate organelle membranes, cellular processes, endoplasmic reticulum, and biological processes in GO analysis.

SRPK1 and EHMT2 collaboratively facilitate the opening of the RNA polymerase I promoter, while WFS1 and C4A are involved in regulating the transport and uptake of Insulin-like Growth Factor (IGF) through Insulin-like Growth Factor Binding Proteins (IGFBPs). Additionally, CCNE2 and ERBB3 may contribute to the development of DR via the GPCR pathway. Both C4A and RAB5B are also linked to the innate immune system, suggesting their involvement in immune-related processes. Furthermore, TP53INP1, CCNE2, and EHMT2 may play a role in DR by regulating TP53 activity, which is critical in cellular stress responses and apoptosis (Supplement Data 39).

# Elevated glycohyocholate and BAFF-R deficiency as protective factors against diabetic retinopathy: insights from metabolomics and immunohistology

Metabolomics research has highlighted numerous bright biomarkers. Glycine-bound bile acids may be protective against both atrophic and neovascular AMD [66]. And there are also clinical studies indicating that elevated glycocholic acid is present in T2D [67]. Our findings advance the research above that elevated Glycohyocholate levels are protective against the development of DR, correlating pre-existing research mechanisms with DR diseases. The rise in hexanoylcarnitine has a contributory effect on the development of DR. In a case-control study, the hexanoylcarnitine concentration was higher than usual in T2D, DR, and NPDR in decreasing order [68].Another metabolic marker of note is 1-palmitoyl-2-oleoyl-GPE (16:0/18:1) levels, which negatively regulate the development of DR. Its link with triglycerides and another 11 methylated genes suggests that methylation may be associated with diabetes risk/complications shared by these genes [69]. To summarise, metabolomics studies have greatly enriched our vision of DR mechanisms, and in addition to the representative molecules mentioned above, many potential molecules deserve to be explored.

Our deep dive into immunohistology showed potential results. Although the corrected results were mostly excluded, we believe that most of these positive results from the primary screening have potential research value. Data from clinical studies suggest a correlation between elevated levels of expression of molecules of the BAFF system and elevated B-cell responses. Although pathogen-mediated increases in ligand and/or receptor expression levels appear to facilitate microbial clearance, certain pathogens have evolved to ablate the B-cell response by inhibiting TACI and/or BAFF-R expression on B cells [70]. Human BAFF-R deficiency is characterized by a paucity of circulating B cells, very low serum IgM and IgG concentrations, but normal or high IgA levels [71]. Our results show a potential causal association between BAFF-R on different B cells and reduced probability of DR. This supports our hypothesis that BAFF-R expression is decreased in patients with DR. More in-depth studies are needed to explore specific BAFF-R-related DR pathogenesis for DR diagnosis and treatment. NKT cells exacerbate retinal white matter stasis and permeability [72]. Our results show that CD16-CD56 on Natural Killer T cells are causally responsible for exacerbating the occurrence of DR. This complements studies that Natural Killer T cells may exacerbate the onset of PDR and suggests a new surface marker: CD16-CD56 [73]. In conclusion, for multiple states of DR, Natural Killer T cells may have a corresponding mediating role. Our study also pointing out that plasmacytoid dendritic cells, plasma blast cell and hematopoietic stem cells are causally associated with a decline in the onset of DR. This further deepens the understanding of the association between these three types of cells and DR [74].

# Strengths and limitations

Our study has multiple strengths. Firstly, we used extensive GWAS data (DR and PDR) and functional enrichment of the transcriptome. The cross-tissue (Pancreas, Kidney, and Whole Blood) gene expression weights created from the tri-tissue data we used increased the statistical power of TWAS, allowing us to mine more genes. Secondly, FUSION post-analysis, such as conditional analysis and permutation testing, helps us identify genes causally related to DR. FOCUS fine-mapping highlights the presence of false-positive genes in our genes. Our cis-eqtl instrument for drug targets defines each target more precisely, allowing us to localize to causally linked genes more accurately. Third, we chose whole-histology weights and targeted three tissues for in-depth exploration (pancreas, renal cortex, and whole blood). Multiple tissues give us a comprehensive and sophisticated view of the tissue of specific interest. Finally, we used PheWAS to explain other multidirectional effects that may have the potential for significant top SNPs. Beyond transcriptomics, we used 731 immune cell databases and 1400 blood metabolite databases for MR analysis on DR. In addition, we validated the knockout mouse expression of the significant genes on the MGI database. Collectively, the comprehensive multi-omics approach of our study expands the horizons of the new wave of diagnosis, prevention, and treatment of DR. We present numerous new insights into drug targets for DR and PDR and present a biological characterization of these targets. Currently, no similar study mines the underlying genes of DR from a drug target multi-omics direction, and our study can provide a valuable contribution to the progress of this part of the research. The molecules in the results of DR play an important role in the early screening and treatment of DR. What's more, the difference between PDR and DR can be applied in curbing the progression of DR and help control the precious rescue and cure time window.

There are several limitations to our study. Firstly, we selected only the cis-eQTL data for analysis to minimize confounding factors; however, trans-eQTL data can also provide valuable insights into gene expression. This limitation may have led to the omission of potentially positive results. Secondly, the demographic source of our data is restricted to European populations, which may hinder the generalization of our findings to Asian and other populations. To address this, access to a more diverse database is needed to complement our study. Our databases for DR and PDR are derived from the Finnish database, raising the possibility of population overlap that could influence our results. Additionally, we did not incorporate databases from other diseases for comparison, which limits our macro-level perspective on disease differentiation and shared mechanisms. Furthermore, while all results from each analysis (e.g., TWAS) are included in Supplementary Data 34, it is important to note that some specific genes may be excluded as false negatives, which could impact our understanding of the factors mediating DR occurrence. To enhance the clinical relevance of our findings, future research should address these limitations by examining potential biases in the data and considering how these insights can inform clinical practice and management of DR and PDR. Researchers can further interpret our study in light of these considerations.

Future research should focus on elucidating the biological mechanisms of RPS26, SRPK1, and WFS1—three critical molecules involved in DR—and the development of drug targets related to them. Our study has established a foundation for advancements in this area. Additionally, there is a need to prioritize early diagnosis and treatment of DR, as well as to analyze its progression from a multiomics perspective.

# Conclusions

In conclusion, our multi-omic transcriptomic analysis of DR and PDR identified key genes, including RPS26, WFS1, and SRPK1, and highlighted differences between DR and PDR at the transcriptomic level, particularly with DCLRE1B. While revealing complex pathogenic factors through metabolomics and immunomics, our study is limited by potential biases from population stratification and environmental confounders. Future research should validate these findings through functional studies and explore diverse populations to better understand the molecular mechanisms underlying DR and PDR. These insights could guide public health strategies for early detection and targeted therapies, reducing diabetic complications and supporting policies for equitable access to advanced treatments.

# Abbreviations

DR	Diabetic retinopathy
PDR	Proliferative diabetic retinopathy
SNPs	Single nucleotide polymorphisms
TWAS	Transcriptome-wide association studies
SMR	Summary data-based mendelian randomization
eQTL	Expression quantitative trait loci
FDR	False discovery rate
PheWAS	Phenome-wide association studies
GWAS	Genome-wide association study
PP	Posterior probability
PIP	Posterior inclusion probability
GO	Gene Ontology
CC	Cellular component
MF	Molecular function
IVW	Inverse variance weighted
scca	Sparse Canonical Correlation Analysis
STROBE	Strengthening the Reporting of Observational Studies in
	Epidemiology
CLSA	Canadian Longitudinal Study of Aging
AC	Absolute cell
MFI	Median fluorescence intensities
RC	Relative cell

LD	Linkage disequilibrium
MGI	Mouse genome informatics
MR-PRESSO	Mendelian Randomization Pleiotropy RESidual Sum and
	Outlier
RPS26	Ribosomal protein S26
WFS1	Wolframin ER transmembrane glycoprotein
SRPK1	SRSF protein kinase 1
BABAM1	BRISC and BRCA1 A complex member 1
RAB5B	Ras-related protein Rab-5B
SENP2	SUMO specific peptidase 2
DCLRE1B	DNA cross-link repair 1B
ITGB7	Integrin subunit beta 7
TEK	Tyrosine-protein kinase receptor
ITPR3	Inositol 1,4,5-trisphosphate receptor type 3
C4A	Complement component 4A
PRKD2	Protein kinase D2
ERBB3	Erb-B2 receptor tyrosine kinase 3
LRRC37A15P	Leucine rich repeat containing 37 member A15, pseudogene
KRT8P46	Keratin 8 pseudogene 46
SUOX	Sulfite oxidase

# **Supplementary Information**

The online version contains supplementary material available at https://doi. org/10.1186/s12967-024-05856-7.

#### Supplementary Material 1.

Supplementary Material 2. Figure 2-1. Heatmap of TWAS analysis results in DR. The threshold is bonferroni correction. It is calculated as follows: *Zcr itical=qnorm*(1–0.05/2*n*). (A)Results of DR in pancreatic tissue. (B)Results of DR in whole blood tissue. (C)Results of DR in kidney cortex tissue. (D) Results of DR in sCCA1. (E)Results of DR in sCCA2. (F)Selection process for significantly positive genes. T: TWAS, T.C: TWAS and colocalization, T.C.P: TWAS, colocalization and permutation testing, T.C.PC: TWAS, colocalization, testing and conditional test, T.C.PC.F: TWAS, colocalization, testing, conditional test and fine-mapping.

Supplementary Material 3. Figure 2-2. Heatmap of TWAS analysis results in PDR. The threshold is bonferroni correction. It is calculated as follows: *Zcri tical=qnorm*(1–0.05/2*n*). (A) Results of PDR in pancreatic tissue. (B) Results of PDR in whole blood tissue. (C)Results of PDR in kidney cortex tissue. (D) Results of PDR in sCCA1. (E) Results of PDR in sCCA2. (F) Selection process for significantly positive genes. T: TWAS, T.C.: TWAS and colocalization, T.C.P: TWAS, colocalization and permutation testing, T.C.PC: TWAS, colocalization testing and conditional test, T.C.PC.F: TWAS, colocalization, permutation testing, conditional test and fine-mapping.

### Acknowledgements

Not applicable.

### Author contributions

YGG, LZR and SYX were the major contributors to the research and the writers of the manuscript. YGG provided major study design. LZR provided technical analysis support and writing work. SYX assisted to visualization and data preparation. FM and CMZ conceived the study, critically reviewed the intellectual content of the manuscript and made substantive revisions to the important contents of the manuscript. MXY, WZJ and ZZR helped revised the manuscript. CZJ provided valuable suggestions. CD and CJK made assistance to the colection of dataset and its analysis. All authors contributed to the article and approved the submitted version.

### Funding

Thanks to the fundings: President's Fund of Zhujiang Hospital of Southern Medical University (ID = yzjj2023ms05); Scientific Research Project of Guangdong Provincial Bureau of Traditional Chinese Medicine (ID = 20241201).

#### Data availability

All analyses in this study were conducted using publicly available data. URLs for the source datasets are as follows: Proliferative diabetic retinopathy: https://storage.googleapis.com/finngen-public-data-r6/summary\_stats/finng

en\_R6\_DM\_RETINA\_PROLIF.gz; Diabetic retinopathy: https://storage.googl eapis.com/finngen-public-data-r9/summary stats/finngen R9 DM RETIN OPATHY\_EXMORE.gc; GTEx\_V8 pancreas weights dataset: https://s3.us-west-1. amazonaws.com/gtex.v8.fusion/ALL/GTExv8.ALL.Pancreas.tar.gz; GTEx\_V8 whole blood weights dataset: https://s3.us-west-1.amazonaws.com/gtex.v8. fusion/ALL/GTExv8.ALL.Whole\_Blood.tar.gz; GTEx\_V8 kidney cortex weights dataset: https://s3.us-west-1.amazonaws.com/gtex.v8.fusion/ALL/GTExv8.ALL. Kidney\_Cortex.tar.gz; sCCA weights dataset: http://gusevlab.org/projects/ fusion/weights/sCCA\_weights\_v8\_2.zip; 1000 Genomes Project Phase 3 European genomic reference data (used for transcriptomic imputation and MR): http://gusevlab.org/projects/fusion/; eQTLgen whole blood eQTL data used for MR of the druggable genome: https://www.eqtlgen.org/; Psychencode dataset: http://development.psychencode.org/; 1400 metabolomics GWAS summary statistics used for MR: https://www.ebi.ac.uk/gwas/, GCST90199621-902010209; Immune cell trait GWAS summary statistics used for MR: https:// gwas.mrcieu.ac.uk/, ebi-a-90001391 through ebi-a-90002121; Dataset used for PheWAS analyses: https://gwas.mrcieu.ac.uk/phewas/.

#### Code availability

The software used in this study are available at the following online repositories. R package TwoSampleMR version 0.5.11: https://github.com/MRCIEU/ TwoSampleMR/releases/tag/v0.5.11; R package ggplot2 version 3.5.1: https:// github.com/tidyverse/ggplot2/releases/tag/v3.5.1; Python package FOCUS version 0.6.10: https://github.com/gusevlab/fusion\_twas; Python package FOCUS (Fine-mapping Of CaUsal gene Sets): https://github.com/bogdanlab/ focus; SMR software: http://cnsgenomics.com/software/smr/\_ R package coloc version 5.2.3: https://cran.r-project.org/web/packages/coloc/index.html; Fig. 1 was made using Office Power Point 2021. Figures 2 and 11 were made using the R package circlize version 0.4.11: https://jokergoo.github.io/circlize/. Figure 3 and 5 were made using R package TWAS Plotter version 1.0: https:// github.com/opain/TWAS-plotter. Figure 4, 6 were made using Python package FOCUS (Fine-mapping Of CaUsal gene Sets): https://github.com/bogda nlab/focus; Fig. 7 were made using R package aPEAR version 1.0.0: https:// cran.r-project.org/web/packages/aPEAR/index.html; Fig. 8, 9 were made using SMR software: https://yanglab.westlake.edu.cn/software/smr/. Figure 10 were made using R package ggplot2 version 3.5.1: https://github.com/tidyverse/ ggplot2/releases/tag/v3.5.1; and locuscomparer version 1.0.0: https://github. com/boxiangliu/locuscomparer.

### Declarations

#### Ethics approval and consent to participate

Our study was a secondary analysis of public dataset. The ethics approval and consent to participate have been obtained from individuals by the original dataset authors.

### **Consent for publication**

Not applicable

### **Competing interests**

There are no conflicts of interest to declare.

### Author details

<sup>1</sup>Department of Ophthalmology, The Sixth Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China. <sup>2</sup>Department of Ophthalmology, Zhujiang Hospital, Southern Medical University, 253 Industrial Avenue, Guangzhou 510282, China. <sup>3</sup>The Second Clinical Medicine School, Southern Medical University, Guangzhou, Guangdong, China. <sup>4</sup>Department of Ophthalmology, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen 518107, Guangdong Province, China. <sup>5</sup>School of Architecture, South China University of Technology, Guangzhou, Guangdong, China. <sup>6</sup>School of Computer Science and Engineering, Sun Yat-Sen University, Guangzhou, China. <sup>7</sup> The Department of Biomedical Innovation Center, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China.

Received: 21 August 2024 Accepted: 2 November 2024 Published online: 24 December 2024

### References

- Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. Lancet. 2010;376:124–36. https://doi.org/10.1016/s0140-6736(09)62124-3.
- Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, et al. Screening for diabetic retinopathy: new perspectives and challenges. Lancet Diabetes Endocrinol. 2020;8:337–47. https://doi.org/10.1016/ s2213-8587(19)30411-5.
- Zhou J, Chen B. Retinal cell damage in diabetic retinopathy. Cells. 2023;12:1342. https://doi.org/10.3390/cells12091342.
- Zheng X, Wan J, Tan G. The mechanisms of NLRP3 inflammasome/pyroptosis activation and their role in diabetic retinopathy. Front Immunol. 2023. https://doi.org/10.3389/fimmu.2023.1151185.
- Lind M, Pivodic A, Svensson A-M, Ólafsdóttir AF, Wedel H, Ludvigsson J. HbA1c level as a risk factor for retinopathy and nephropathy in children and adults with type 1 diabetes: Swedish population based cohort study. BMJ. 2019. https://doi.org/10.1136/bmj.l4894.
- Perais J, Agarwal R, Evans JR, Loveman E, Colquitt JL, Owens D, et al. Prognostic factors for the development and progression of proliferative diabetic retinopathy in people with diabetic retinopathy. Cochrane Libr. 2023. https://doi.org/10.1002/14651858.cd013775.pub2.
- Nawaz IM, Rezzola S, Cancarini A, Russo A, Costagliola C, Semeraro F, et al. Human vitreous in proliferative diabetic retinopathy: characterization and translational implications. Prog Retin Eye Res. 2019;72: 100756. https:// doi.org/10.1016/j.preteyeres.2019.03.002.
- Sood A, Baishnab S, Gautam I, Choudhary P, Lang DK, Jaura RS, et al. Exploring various novel diagnostic and therapeutic approaches in treating diabetic retinopathy. Inflammopharmacology. 2023;31:773–86. https://doi.org/10.1007/s10787-023-01143-x.
- Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613:508–18. https://doi.org/10.1038/ s41586-022-05473-8.
- Feng H, Mancuso N, Gusev A, Majumdar A, Major M, Pasaniuc B, et al. Leveraging expression from multiple tissues using sparse canonical correlation analysis and aggregate tests improves the power of transcriptomewide association studies. PLoS Genet. 2021;17: e1008973. https://doi.org/ 10.1371/journal.pgen.1008973.
- The GTEx Consortium, Aguet F, Anand S, Ardlie KG, Gabriel S, Getz GA, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369:1318–30. https://doi.org/10.1126/scien ce.aaz1776.
- Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016;48:245–52. https://doi.org/10.1038/ng.3506.
- Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10: e1004383. https://doi.org/10.1371/journal.pgen.1004383.
- Zhao S, Crouse W, Qian S, Luo K, Stephens M, He X. Adjusting for genetic confounders in transcriptome-wide association studies improves discovery of risk genes of complex traits. Nat Genet. 2024;56:336–47. https://doi. org/10.1038/s41588-023-01648-9.
- Mancuso N, Freund MK, Johnson R, Shi H, Kichaev G, Gusev A, et al. Probabilistic fine-mapping of transcriptome-wide association studies. Nat Genet. 2019;51:675–82. https://doi.org/10.1038/s41588-019-0367-1.
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb). 2021;2: 100141. https://doi.org/10.1016/j.xinn.2021.100141.
- Yu G, Wang L-G, Han Y, He Q-Y. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7. https://doi.org/10.1089/omi.2011.0118.
- Kerseviciute I, Gordevicius J. aPEAR:an R package for autonomous visualisation of pathway enrichment networks. bioRxiv. 2023;
- Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48:481–7. https://doi.org/10.1038/ng.3538.
- Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53:1300–10. https://doi.org/10.1038/s41588-021-00913-z.

- Li M, Santpere G, Imamura Kawasawa Y, Evgrafov OV, Gulden FO, Pochareddy S, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science. 2018. https://doi. org/10.1126/science.aat7615.
- Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. Stat Methods Med Res. 2012;21:223–42. https://doi.org/ 10.1177/0962280210394459.
- Gill D, Efstathiadou A, Cawood K, Tzoulaki I, Dehghan A. Education protects against coronary heart disease and stroke independently of cognitive function: evidence from Mendelian randomization. Int J Epidemiol. 2019;48:1468–77. https://doi.org/10.1093/ije/dyz200.
- EPIC-InterAct Consortium, Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol. 2015;30:543–52. https://doi.org/10.1007/s10654-015-0011-z.
- Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512–25. https://doi.org/10.1093/ije/ dyv080.
- 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. https:// doi.org/10.1002/gepi.21965.
- 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. https://doi.org/10.1093/ije/dyx102.
- Lin J, Zhou J, Xu Y. Potential drug targets for multiple sclerosis identified through Mendelian randomization analysis. Brain. 2023;146:3364–72. https://doi.org/10.1093/brain/awad070.
- Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the reporting of Observational Studies in Epidemiology using Mendelian randomization: The STROBE-MR Statement: The STROBE-MR statement. JAMA [Internet]. 2021;326:1614–21. Available from: https://doi.org/10.1001/jama.2021.18236.
- Bastarache L, Denny JC, Roden DM. Phenome-wide association studies. JAMA. 2022;327:75. https://doi.org/10.1001/jama.2021.20356.
- Chen Y, Lu T, Pettersson-Kymmer U, Stewart ID, Butler-Laporte G, Nakanishi T, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. Nat Genet. 2023;55:44–53. https://doi. org/10.1038/s41588-022-01270-1.
- Orrù V, Steri M, Sidore C, Marongiu M, Serra V, Olla S, et al. Complex genetic signatures in immune cells underlie autoimmunity and inform therapy. Nat Genet. 2020;52:1036–45. https://doi.org/10.1038/ s41588-020-0684-4.
- Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50:693–8. https:// doi.org/10.1038/s41588-018-0099-7.
- MGI-Mouse Genome Informatics-The international database resource for the laboratory mouse. Org. http://www.informatics.jax.org. Accessed 21 Aug 2021.
- Peinado H, Alečković M, Lavotshkin S, Matei I, Costa-Silva B, Moreno-Bueno G, et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. Nat Med. 2012;18:883–91. https://doi.org/10.1038/nm.2753.
- Farrar JE, Vlachos A, Atsidaftos E, Carlson-Donohoe H, Markello TC, Arceci RJ, et al. Ribosomal protein gene deletions in Diamond-Blackfan anemia. Blood. 2011;118:6943–51. https://doi.org/10.1182/blood-2011-08-375170.
- Sahay M, Kalra S, Badani R, Bantwal G, Bhoraskar A, Das AK, et al. Diabetes and anemia: International Diabetes Federation (IDF)—Southeast Asian Region (SEAR) position statement. Diabetes Metab Syndr. 2017;11:S685– 95. https://doi.org/10.1016/j.dsx.2017.04.026.
- Li Y, Xu J, Lu Y, Bian H, Yang L, Wu H, et al. DRAK2 aggravates nonalcoholic fatty liver disease progression through SRSF6-associated RNA alternative splicing. Cell Metab. 2021;33:2004-2020.e9. https://doi.org/10.1016/j. cmet.2021.09.008.
- Yang Y-M, Jung Y, Abegg D, Adibekian A, Carroll KS, Karbstein K. Chaperone-directed ribosome repair after oxidative damage. Mol Cell. 2023;83:1527-1537.e5. https://doi.org/10.1016/j.molcel.2023.03.030.

- Yang Y-M, Karbstein K. The chaperone Tsr2 regulates Rps26 release and reincorporation from mature ribosomes to enable a reversible, ribosomemediated response to stress. Sci Adv. 2022. https://doi.org/10.1126/ sciadv.abl4386.
- Richardson TG, Hemani G, Gaunt TR, Relton CL, Davey SG. A transcriptome-wide Mendelian randomization study to uncover tissue-dependent regulatory mechanisms across the human phenome. Nat Commun. 2020. https://doi.org/10.1038/s41467-019-13921-9.
- Chapla A, Johnson J, Korula S, Mohan N, Ahmed A, Varghese D, et al. WFS1 gene-associated diabetes phenotypes and identification of a founder mutation in Southern India. J Clin Endocrinol Metab. 2022;107:1328–36. https://doi.org/10.1210/clinem/dgac002.
- Wang L, Liu H, Zhang X, Song E, Wang Y, Xu T, et al. WFS1 functions in ER export of vesicular cargo proteins in pancreatic β-cells. Nat Commun. 2021. https://doi.org/10.1038/s41467-021-27344-y.
- 44. Köks S. Genomics of wolfram Syndrome 1 (WFS1). Biomolecules. 2023. https://doi.org/10.3390/biom13091346.
- Liu H, Gong Z, Li K, Zhang Q, Xu Z, Xu Y. SRPK1/2 and PP1α exert opposite functions by modulating SRSF1-guided MKNK2 alternative splicing in colon adenocarcinoma. J Exp Clin Cancer Res. 2021. https://doi.org/10. 1186/s13046-021-01877-y.
- Gammons MVR, Dick AD, Harper SJ, Bates DO. SRPK1 inhibition modulates VEGF splicing to reduce pathological neovascularization in a rat model of retinopathy of prematurity. Invest Ophthalmol Vis Sci. 2013;54:5797. https://doi.org/10.1167/iovs.13-11634.
- Amin EM, Oltean S, Hua J, Gammons MVR, Hamdollah-Zadeh M, Welsh GI, et al. WT1 mutants reveal SRPK1 to be a downstream angiogenesis target by altering VEGF splicing. Cancer Cell. 2011;20:768–80. https://doi.org/10. 1016/j.ccr.2011.10.016.
- Malhi NK, Allen CL, Stewart E, Horton KL, Riu F, Batson J, et al. Serinearginine-rich protein kinase-1 inhibition for the treatment of diabetic retinopathy. Am J Physiol Heart Circ Physiol. 2022;322:H1014–27. https:// doi.org/10.1152/ajpheart.00001.2022.
- Gammons MV, Fedorov O, Ivison D, Du C, Clark T, Hopkins C, et al. Topical antiangiogenic SRPK1 inhibitors reduce choroidal neovascularization in rodent models of exudative AMD. Invest Ophthalmol Vis Sci. 2013;54:6052. https://doi.org/10.1167/iovs.13-12422.
- Xu Q, Liu X, Liu Z, Zhou Z, Wang Y, Tu J, et al. MicroRNA-1296 inhibits metastasis and epithelial-mesenchymal transition of hepatocellular carcinoma by targeting SRPK1-mediated PI3K/AKT pathway. Mol Cancer. 2017. https://doi.org/10.1186/s12943-017-0675-y.
- Zhao Y, Wu D, Jiang D, Zhang X, Wu T, Cui J, et al. A sequential methodology for the rapid identification and characterization of breast cancerassociated functional SNPs. Nat Commun. 2020. https://doi.org/10.1038/ s41467-020-17159-8.
- Feng L, Huang J, Chen J. MERIT40 facilitates BRCA1 localization and DNA damage repair. Genes Dev. 2009;23:719–28. https://doi.org/10.1101/gad. 1770609.
- Wang F-W, Cao C-H, Han K, Zhao Y-X, Cai M-Y, Xiang Z-C, et al. APC-activated long noncoding RNA inhibits colorectal carcinoma pathogenesis through reduction of exosome production. J Clin Invest. 2021. https://doi. org/10.1172/jci149666.
- He X, Kuang G, Wu Y, Ou C. Emerging roles of exosomal miRNAs in diabetes mellitus. Clin Transl Med. 2021. https://doi.org/10.1002/ctm2.468.
- Heo K-S, Chang E, Le N-T, Cushman H, Yeh ETH, Fujiwara K, et al. De-SUMOylation enzyme of sentrin/SUMO-specific protease 2 regulates disturbed flow–induced SUMOylation of ERK5 and p53 that leads to endothelial dysfunction and atherosclerosis. Circ Res. 2013;112:911–23. https://doi.org/10.1161/circresaha.111.300179.
- Atla G, Bonàs-Guarch S, Cuenca-Ardura M, Beucher A, Crouch DJM, Garcia-Hurtado J, et al. Genetic regulation of RNA splicing in human pancreatic islets. Genome Biol. 2022. https://doi.org/10.1186/ s13059-022-02757-0.
- Baddock HT, Newman JA, Yosaatmadja Y, Bielinski M, Schofield CJ, Gileadi O, et al. A phosphate binding pocket is a key determinant of exo- versus endo-nucleolytic activity in the SNM1 nuclease family. Nucleic Acids Res. 2021;49:9294–309. https://doi.org/10.1093/nar/gkab692.
- Sonmez C, Toia B, Eickhoff P, Matei AM, El Beyrouthy M, Wallner B, et al. DNA-PK controls Apollo's access to leading-end telomeres. Nucleic Acids Res. 2024;52:4313–27. https://doi.org/10.1093/nar/gkae105.

- 59. Xiao Y, Wang C, Chen J-Y, Lu F, Wang J, Hou N, et al. Deficiency of PRKD2 triggers hyperinsulinemia and metabolic disorders. Nat Commun. 2018. https://doi.org/10.1038/s41467-018-04352-z.
- Yellowlees Douglas J, Bhatwadekar AD, Li Calzi S, Shaw LC, Carnegie D, Caballero S, et al. Bone marrow-CNS connections: implications in the pathogenesis of diabetic retinopathy. Prog Retin Eye Res. 2012;31:481–94. https://doi.org/10.1016/j.preteyeres.2012.04.005.
- Azoitei N, Diepold K, Brunner C, Rouhi A, Genze F, Becher A, et al. HSP90 supports tumor growth and angiogenesis through PRKD2 protein stabilization. Cancer Res. 2014;74:7125–36. https://doi.org/10.1158/0008-5472. can-14-1017.
- 62. Fan Y, Murgia M, Linder MI, Mizoguchi Y, Wang C, Łyszkiewicz M, et al. HAX1-dependent control of mitochondrial proteostasis governs neutrophil granulocyte differentiation. J Clin Invest. 2022. https://doi.org/10. 1172/jci153153.
- Wu L, Islam MR, Lee J, Takase H, Guo S, Andrews AM, et al. ErbB3 is a critical regulator of cytoskeletal dynamics in brain microvascular endothelial cells: Implications for vascular remodeling and blood brain barrier modulation. J Cereb Blood Flow Metab. 2021;41:2242–55. https://doi.org/10.1177/0271678x20984976.
- Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An expanded genome-wide association study of type 2 diabetes in Europeans. Diabetes. 2017;66:2888–902. https://doi.org/10.2337/db16-1253.
- Onengut-Gumuscu S, Chen W-M, Robertson CC, Bonnie JK, Farber E, Zhu Z, et al. Type 1 diabetes risk in African-ancestry participants and utility of an ancestry-specific genetic risk score. Diabetes Care. 2019;42:406–15. https://doi.org/10.2337/dc18-1727.
- Warden C, Brantley MA Jr. Glycine-conjugated bile acids protect RPE tight junctions against oxidative stress and inhibit choroidal endothelial cell angiogenesis in vitro. Biomolecules. 2021;11:626. https://doi.org/10.3390/ biom11050626.
- Banimfreg BH, Shamayleh A, Alshraideh H, Semreen MH, Soares NC. Untargeted approach to investigating the metabolomics profile of type 2 diabetes emiratis. J Proteomics. 2022;269: 104718. https://doi.org/10. 1016/j.jprot.2022.104718.
- Wang Z, Tang J, Jin E, Ren C, Li S, Zhang L, et al. Metabolomic comparison followed by cross-validation of enzyme-linked immunosorbent assay to reveal potential biomarkers of diabetic retinopathy in Chinese with type 2 diabetes. Front Endocrinol (Lausanne). 2022. https://doi.org/10.3389/ fendo.2022.986303.
- 69. Yousri NA, Albagha OME, Hunt SC. Integrated epigenome, whole genome sequence and metabolome analyses identify novel multi-omics pathways in type 2 diabetes: a Middle Eastern study. BMC Med. 2023. https://doi.org/10.1186/s12916-023-03027-x.
- Sakai J, Akkoyunlu M. The role of BAFF system molecules in host response to pathogens. Clin Microbiol Rev. 2017;30:991–1014. https://doi.org/10. 1128/cmr.00046-17.
- Smulski CR, Eibel H. BAFF and BAFF-receptor in B cell selection and survival. Front Immunol. 2018. https://doi.org/10.3389/fimmu.2018.02285.
- Suvas P, Liu L, Rao P, Steinle JJ, Suvas S. Systemic alterations in leukocyte subsets and the protective role of NKT cells in the mouse model of diabetic retinopathy. Exp Eye Res. 2020;200: 108203. https://doi.org/10. 1016/j.exer.2020.108203.
- Agardh E, Lundstig A, Perfilyev A, Volkov P, Freiburghaus T, Lindholm E, et al. Genome-wide analysis of DNA methylation in subjects with type 1 diabetes identifies epigenetic modifications associated with proliferative diabetic retinopathy. BMC Med. 2015. https://doi.org/10.1186/ s12916-015-0421-5.
- Amato R, Catalani E, Dal Monte M, Cammalleri M, Cervia D, Casini G. Morpho-functional analysis of the early changes induced in retinal ganglion cells by the onset of diabetic retinopathy: the effects of a neuroprotective strategy. Pharmacol Res. 2022;185: 106516. https://doi.org/10.1016/j.phrs. 2022.106516.

# Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.