Diabetes, Metabolic Syndrome and Obesity

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

CD44, ACAN, PLVAP, and HBEGF Emerged as Potential Biomarkers in Diabetic Retinopathy

Qingbo Li[1,*], Xu Zhou[1,*], Ying Wang[2,*], Yi Peng[1], Juping Liu[1], Xiaorong Li[1], Yan Shao[1,2]

[1]Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital, Tianjin, 300384, People's Republic of China; [2]Modern Medical Teaching and Research Section, Department of Tibetan Medicine, University of Tibetan Medicine, Lhasa, Tibet Autonomous Region, 850000, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yan Shao, Email sytmueh@163.com

Background: DNA methylation was critically involved in the occurrence and development of diabetic retinopathy (DR), but the clinical diagnostic value and mechanism remains unclear. This study aimed to identify and validate the biomarkers associated with DNA methylation in DR.

Methods: Public databases provided the data used in this study. The biomarkers associated with DNA methylation in proliferative diabetic retinopathy (PDR) and non-proliferative diabetic retinopathy (NPDR) were identified by differential expression analysis, protein-protein interaction network construction, mendelian randomization, and 2 machine learning algorithms, respectively. Gene set enrichment analysis was utilized to determine the functional roles of these biomarkers. The cell-specific expression patterns of biomarkers were analyzed at the single-cell level. Immune infiltration analysis was performed to further study the relationship between biomarkers, and 28 types of immune cell infiltration. The diagnostic power of biomarkers was assessed by the receiver operating curve (ROC) and nomogram.

Results: CD44 was preliminarily identified as a biomarker linked to PDR, whereas ACAN, PLVAP, and HBEGF emerged as biomarkers associated with NPDR. The biomarkers were involved ibosome, lysosome, oxidative phosphorylation, aminoacyl tRNA biosynthesis, P53 signaling pathway, and other pathways. In the retina of diabetic mice, PLVAP and HBEGF were mainly enriched in neutrophils. In the retinal fibroproliferative membranes of PDR patients, CD44 were mainly enriched in monocytes/macrophages. Moreover, CD44 was positively correlated with the remaining differential immune cells except for type 2 T helper cells. Effector memory CD4 T cells showed the strongest positive correlation with ACAN and the largest negative correlation with HBEGF. Finally, the ROC and nomogram validated biomarkers had an excellent diagnostic efficacy.

Conclusion: This study identified 4 genes (CD44, ACAN, PLVAP, and HBEGF) associated with DNA methylation may serve as biomarkers of DR, offering new insights into potential therapeutic strategies for DR.

Keywords: diabetic retinopathy, DNA methylation, transcriptome, single-cell RNA sequencing analysis, Mendelian randomization, machine learning

Introduction

The International Diabetes Federation reported that diabetes mellitus affects approximately 536.6 million adults worldwide, with projections suggesting an increase to 783 million by 2045.[1] A common microvascular complication of diabetes is diabetic retinopathy (DR), which is one of the leading causes of acquired blindness among adults aged 50 years and older.[2] Based on the presence or absence of retinal neovascularization, DR is clinically classified into 2 stages: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). Although intravitreal anti-vascular endothelial growth factor agents have revolutionized the treatment of DR,[3] the underlying retinal ischemia is unchanged and the efficacy can diminish over time.[4,5] Therefore, there is an urgent need to identify novel biomarkers and target-specific therapies.

Received: 9 August 2025
Accepted: 19 September 2025
Published: 3 October 2025

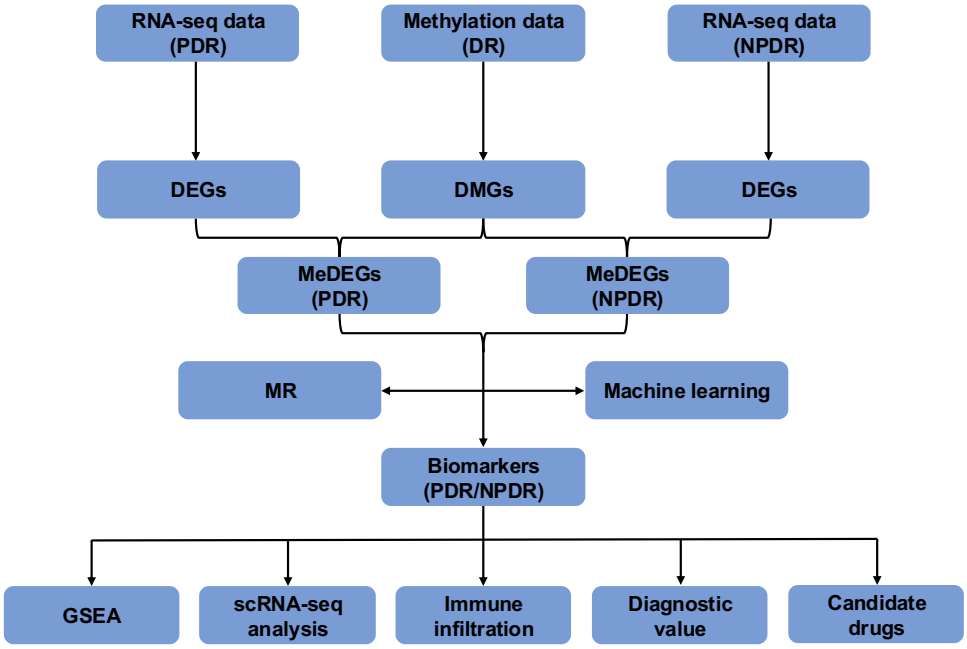

Figure 1 Data analysis flowchart.
Abbreviations: DR, diabetic retinopathy; NPDR, non-proliferative diabetic retinopathy; PDR, proliferative diabetic retinopathy; RNA-seq, RNA-sequencing; DEGs, Differentially expressed genes; DMGs, differentially methylated genes; MeDEGs, methylation-regulated differentially expressed genes; MR, mendelian randomization; GSEA, gene set enrichment analysis; scRNA-seq, Single-cell RNA sequencing.

DNA methylation is a common post-replication epigenetic modification in both prokaryotic and eukaryotic genomes, regulating genome function without altering the primary structure of the DNA molecule.[6] This process is mediated by the transfer of a methyl group from S-adenosylmethionine to a cytosine at CpG dinucleotide sites, catalyzed by DNA methyltransferase.[7] Aberrant DNA methylation is closely associated with DR and plays a critical role in its occurrence and progression via mechanisms such as oxidative stress, inflammation, neovascularization, metabolic memory, and other complex pathophysiological processes.[8,9] Methylation-regulated differentially expressed genes (MeDEGs) have the potential to serve as biomarkers for diagnosis and early detection of DR.

The Infinium DNA methylation BeadChip platform enables accurate and precise methylation measurements across the human genome.[10] RNA sequencing (RNA-seq) technology has become an essential tool for analyzing differential gene expression at the transcriptome level. Single-cell RNA sequencing (scRNA-seq) technology Single-cell RNA sequencing (scRNA-seq) technology allows for the detection of gene expression at the single cell level, providing deeper insights into the different types of cells within tissues.[11] Although single-omics analysis can provide partial insights, multi-omics analysis provides a more comprehensive understanding of the pathogenesis of DR.[12,13]

In the present study, we integrated methylation array and RNA-seq to identify MeDEGs of DR and screened the biomarkers by machine learning combined with mendelian randomization (MR). Subsequently, we evaluated the cellular localization of biomarkers by integrated scRNA-seq and the relationship between immune cell infiltrates and biomarkers. Finally, we validated the diagnostic value of biomarkers and predicted the corresponding candidate drugs. The analysis process was shown in Figure 1. Our study will be of profound implications for clarifying the role of methylation in DR and identifying the underlying biomarker for DR future research.

Materials and Methods

Data Collection

The Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) provided the PDR RNA-seq datasets GSE102485 (platform: GPL18573) and GSE94019 (platform: GPL11154). GSE102485 contained 27 PDR and 3 normal retinal tissue samples, and GSE94019 included 9 PDR and 4 normal retinal tissue samples. To reduce the possibility of

batch effects due to abiotic bias between datasets, GSE102485 and GSE94019 were integrated into a new total PDR dataset by the Combat_seq algorithm of the sva package (v 3.50.0) (https://bioconductor.org/packages/sva). The total PDR dataset included 36 PDR and 7 normal retinal tissue samples. The RNA-seq dataset GSE160306 (platform: GPL20301) related to NPDR was from the GEO database and contained 19 non-PDR (NPDR) and 20 normal tissues samples. The DNA methylation dataset GSE57362 (platform: GPL13534) was downloaded from the GEO database, and 9 PDR, 8 normal, and 8 NPDR tissue samples were used for this study. The scRNA-seq datasets (GSE165784 and GSE205123) were both obtained from the GEO database. GSE165784 (platform: GPL20795) included 5 PDR and 1 normal fibrous membrane tissue samples (human samples). Meanwhile, GSE205123 (platform: GPL24247) consisted of 1 DR and 1 normal retinal tissue samples (mouse samples). Furthermore, GWAS information on DR (GWAS ID: finngen_R9_DM_RETINOPATHY_EXMORE) was gathered from the IEU Open GWAS (https://gwas.mrcieu.ac.uk/) database, which included 16,380,347 single nucleotide polymorphisms (SNPs) and 190,594 European samples (14,584 DR and 176,010 normal samples). The IEU Open GWAS database provided the exposure factors' expression quantitative trait locus (eQTL) data.

DNA Methylation Analysis

First of all, the DNA methylation data in GSE57362 was imported through the champ function, and the filterDetP function was used to filter the probes with lower read confidence (filtering probes with $P > 0.05$). Then, quality control (QC) was performed using the champ. QC function. Also, noob and Beta-Mixture Quantile (BMIQ) were used for noise reduction and normalization processes, respectively. The information about the density distribution of the probe beta values was displayed. After that, the samples were clustered based on the beta values of the probes.

$$\text{Beta} = \text{M}/(\text{M} + \text{U} + \text{Offest})$$

where M represented the signal intensity of methylation at CpG sites, U represented the signal intensity of unmethylation at CpG sites, and Offset was set to 100.

Subsequently, differential methylation probe (DMP) and differential methylation region (DMR) analyses were performed. Specifically, the limma package (v 3.54.0) was used for differential methylation analysis, and the Benjamini-Hochberg (BH) method was performed to correct for multiple construction tests, and probes were considered to be DMP when they met $P < 0.05$. The champ.DMR function was then used for DMR analysis via the Bumphunter algorithm. To control the false positive rate, minProbes ≥ 10 and adjPvalDmr < 0.05 were used to filter the results. Differentially methylated genes (DMGs) were defined as those that satisfy $|\log_2\text{FoldChange (FC)}| \geq 0.1$ and $P < 0.05$, and whose $\log_2\text{FC} > 0$ for hypermethylated sites and $\log_2\text{FC} < 0$ for hypomethylated sites.

Differential Expression Analysis

In order to view the distribution of PDR and normal samples in the total PDR RNA-seq dataset, the prcomp function of the stats package (v 4.3.1) (https://search.r-project.org/R/refmans/stats/html/00Index.html) was used to perform principal component analysis (PCA). Afterwards, to identify differentially expressed genes 1 (DEGs) between PDR and normal samples in the total PDR RNA-seq dataset, the DESeq2 package (v 1.42.0) was used for differential expression analysis (Benjamini-Hochberg–adjusted $P < 0.05$ and $|\log_2\text{FC}| > 1.5$). In addition, the same analysis was performed for GSE160306 to obtain DEGs in the NPDR RNA-seq dataset. Subsequently, volcano plots were plotted by the ggplot2 package (v 3.4.4), respectively, to visualize the DEGs.

PPI Network Construction

The genes intersecting DMGs with DEGs were taken as MeDEGs, respectively. The ggVennDiagram package (v 1.2.2) was used to draw the Venn diagram. Subsequently, the TRING database (http://www/string-db.org/) was used to construct the PPI network based on PDR MeDEGs. Then, the MNC, degree, and EPC algorithms of Cutohubba plug-in in Cytoscape software (v 3.9.1) were used to obtain the top 50 genes, respectively, and the genes obtained from the intersection of top 50 genes in 3 algorithms were used as intersection genes.

MR Analysis

The intersection genes were used as exposure factors and PDR as the result in the MR study. First, exposure factor reading and instrumental variable (IV) screening were done using the VariantAnnotation software (v 1.71). The most substantially related IVs with the outcome were screened using the $P < 5\times10^{-8}$. The $r^2 = 0.001$, clump = TRUE, and kb = 100 were used to exclude IVs with linkage disequilibrium (LD). In addition, IVs significantly associated with outcome and with F-statistics less than 10 were also excluded. Exposure factors with SNPs less than 3 were also excluded. Further analysis was conducted using the TwoSampleMR program (v 0.5.6). The mr function then carried out MR analysis in conjunction with 5 algorithms: inverse variance weighted (IVW), weighted median, simple mode, MR Egger, as well as weighted mode. Among them, the results of IVW method were used as the main reference. The P-value < 0.05 indicated that there was a significant causal relationship between intersection genes and PDR. After that, to determine the correlation between exposure factors and outcome, scatter plots by mr_scatter_plot function, forest plots by mr_forest_plot function, and funnel plots by mr_funnel_plot function were plotted respectively. In addition, to determine the reliability of the MR results, sensitivity analysis was performed. In particular, the mr_heterogeneity function (P > 0.05) was utilized to conduct the heterogeneity test. The mr _pleiotropy_test function was applied to perform the horizontal pleiotropy test (P > 0.05). In the meantime, the mr_leaveoneout function was used to execute leave-one-out (LOO). The directionality of the causal link between exposure factors and outcomes was evaluated using the MR Steiger test. Lastly, genes that showed a strong and consistent causal link to PDR in MR analysis were deemed key genes.

Identification of the Biomarkers by Machine Learning Algorithms

Based on the key genes screened by MR analysis, in the total PDR dataset, the LASSO using 10-fold cross-validation and the SVM-RFE machine learning algorithms were carried out through glmnet package (v 4.1.8) and caret package (v 6.0.94), respectively. Biomarker of PDR was obtained by taking the intersection of the feature genes derived from these 2 machine learning algorithms through the ggVennDiagram package (v 1.2.2). Meanwhile, the same analysis was performed based on the MeDEGs to identify the biomarkers of NPDR in GSE160306.

Gene Set Enrichment Analysis (GSEA)

Using the psych package (v 2.2.9), Spearman correlation analysis was carried out for biomarkers to better understand the biological processes and pathways involved in biomarkers throughout DR development. Genes were ranked according to correlation coefficient from highest to lowest. The background gene set "c2.cp.kegg.v7.5.1.symbols.gmt" was obtained from the MSigDB (http://www.broadinstitute.org/msigdb). Subsequently, the GSEA function of the clusterProfiler package (v 4.7.1) was used to execute GSEA (P < 0.05, |NES| > 1).

Competing Endogenous RNAs (ceRNAs) Network Analysis

In order to reveal the regulatory mechanisms of biomarkers involved in DR occurrence, miRNAs were predicted through TargetScan database (http://www.targetscan.org). For the miRNAs exceeding 5, the top 5 miRNAs with higher confidence were screened. Then, based on miRNAs, lncRNAs were predicted through the Starbase database (http://starbase.sysu.edu.cn/) (clipExpNum > 10). Subsequently, a lncRNA-miRNA-mRNA regulatory network was constructed.

The scRNA-Seq Analysis

The scRNA-seq analysis was performed in GSE165784 and GSE205123 using Seurat package (v 5.1.0). First, the following criteria were used to filter low-quality cells and genes: 1) cells with nFeature RNA less than 400 and greater than 5000 (GSE165784) or 3000 (GSE205123); 2) cells with nCount RNA less than 1000; 3) cells with mitochondrial percent_mt greater than 20%; and 4) erythrocytes with greater than 5% RNA. Subsequently, the SCTransform function of the Seura package (v 5.1.0) was utilized to determine the top 2000 highly variable genes with large variations in expression levels, labelling the top 10 most variable genes using the LabelPoints package (v 1.71). After that, the data were normalized with the ScaleData function. Then, PCA was undertaken on top 2000 highly variable genes. ElbowPlot function was applied to generate a scree plot. Statistically significant PCs were selected based on the JackStrawPlot function and the scree plot. Next, UMAP was performed (resolution = 0.2) to identify distinct cell clusters through

FindNeighbors function and FindClusters function of Seurat package (v 5.1.0) and k-nearest neighbors (KNN) algorithm. Based on the ScMayoMap function, the cell clusters were annotated by selecting the organization as "eye" to get different cell types. Subsequently, the Wilcoxon test was employed to compare the expression of biomarkers in different cell types between DR/PDR and normal samples ($P < 0.05$). Cell types with more than half of differential expression of biomarkers were considered as key cell.

Immune Infiltration Analysis

Using the ssGSEA algorithm, the infiltration abundance of 28 immune cells in DR samples was evaluated. The Wilcoxon test was then used to evaluate the immune cell infiltration levels between PDR and normal samples and between NPDR and normal samples, respectively ($P < 0.05$). Box plots were created for visualization using the ggplot2 package (v 3.4.4). The psych package (v 2.2.9) was then applied to assess Spearman correlations between various immune cells as well as between biomarkers and differential immune cells ($|cor| > 0.3$, $P < 0.05$).

Construction and Evaluation of Nomogram

The nomograms in the NPDR dataset were created using the rms program (v 6.5–0) to investigate the diagnostic potential of biomarkers. The calibration curve was then created using the calibrate function of the rms package (v 6.5–0) to evaluate the nomogram model's predicted accuracy. The pROC package (version 1.18.5) was used to create the ROC curve. An area under the curve (AUC) value greater than 0.7 indicated good predictive ability of the nomogram model. In addition, decision curve was plotted through the rmda package (v 1.6).

Human Sample Collection

The participants were hospitalized patients with PDR, macular hole between January 2023 and January 2024 at the Tianjin Medical University Eye Hospital in Tianjin, China. PDR was based on the American Academy of Ophthalmology criteria (2019). Exclusion criteria were severe diabetic complications, severe and chronic infection symptoms, type 1 diabetes mellitus, liver and kidney dysfunction, tumors, cardiovascular diseases, and patients with eye diseases other than cataracts or macular hole. Eventually, 19 patients with PDR, 20 with macular hole were included. Informed consent was obtained from all participants prior to the survey, and the Research Ethics Committee of Tianjin Medical University Eye Hospital approved the study (2017KY-01).

DM Mouse Model Construction

Eight-week-old male C57BL/6J mice (SiPiFu, Beijing, China) were housed within ventilated cages in a pathogen-free room at a constant temperature (23±2°C) under a 12-hour light-dark cycle at the Tianjin Medical University Eye Institute Laboratory Animal Center (Tianjin, China). The T1DM mice (n=6) were given a normal diet (10 kcal% fat, D12450J, Research Diets, Keaoxieli, Beijing, China) for 14 days and were then induced by intraperitoneal injection of streptozotocin (STZ, 150 mg/kg; Sigma-Aldrich, St. Louis, MO, USA). T2DM mice (n=6) were fed a high-fat diet (HFD) (60 kcal % fat, D12492, Research Diets, SaiNo, Jilin, China) for 14 days and were then induced with STZ (85 mg/kg), followed by maintaining on a HFD. The normal group (n=6) were given a normal diet and injected with 0.1 M sodium citrate buffer (Solarbio, Beijing, China). Each experimental group contained 6 mice. One week later, mice with plasma glucose levels >16.7 mmol/L were classified as diabetic. All mice were euthanized under isoflurane anesthesia 12 weeks after the plasma glucose levels >16.7 mmol/L.

On postnatal day 7 (P7), the mice were placed in a closed oxygen box with an oxygen concentration of 75%±2%. The mice were returned to room air (oxygen concentration 21%) at P12 and maintained in room air from P12 to P17. Age-matched mice raised in room air were used as the controls. Each experimental group contained 6 mice. All mice were euthanized under isoflurane anesthesia.

The use of animals in the experimental protocol was approved by the Tianjin Medical University Eye Institute Laboratory Animal Center (TJYY20231202894).

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Total cellular RNA was extracted using an RNA extraction kit (EZBioscience, US). Subsequently, cDNA was generated through reverse transcription of RNA, employing the color reverse transcription reagents (EZBioscience, US). For the mouse retina, the primer sequences used were (5′-3′): β-actin (forward: 5′-TAAGGCCAACCGTGAAAAG-3′, reverse: 5′-ACCAGAGGCATACAGGGACA-3′), CD44 (forward: 5′-CGGAACCACAGCCTCCTTTCAA-3′, reverse: 5′-TGCCATCCGTTCTGAAACCACG-3′), ACAN (forward: 5′-CAGGCTATGAGCAGTGTGATGC-3′, reverse: 5′-GCTGCTGTCTTTGTCACCCACA-3′), PLVAP (forward: 5′-ACCCGTGAGAATGCAGAAC-3′, reverse: 5′-TGTCCAGCGCGCTAATG-3′), HBEGF (forward: 5′-GAGTTCCGTACTCCCTCTTGCA-3′, reverse: 5′-CAGCCAAGACTGTAGTGTGGTC-3′) designed using the Primer-BLAST. The primer sequence was synthesized by Shanghai Sangon Biotechnology (Shanghai, China). Calculation of the expression of the target gene was performed based on the $2^{-\Delta\Delta Ct}$ method.

Enzyme-linked Immunosorbent Assay (ELISA)

Human CD44 ELISA kit was obtained from CUSABIO (Wuhan, China). Briefly, 100 μL of sample diluent, prepared standard and vitreous humor samples were added to appropriate wells. The resulting plate was incubated for 2 h at 37°C with plate lid. After removing all liquid completely, 100 μL of biotin-antibody was added to each well. The resulting plate was incubated for 1 h at 37°C with plate lid and subsequently washed three times with wash buffer. After washing the wells, 100 μL of HRP-avidin was added and incubated for 1 h at 37°C and subsequently washed five times with wash buffer. After washing the wells, 90 μL of TMB substrate was added and incubated for 30 min at 37°C. Finally, the reaction was stopped by adding 50 μL stop solution, and the signal was measured at 450 nm.

Statistical Analysis

In order to conduct bioinformatics analysis, R (version 4.2.2) was used. Results are represented as means ± standard deviation. The differences between the 2 groups were compared using the Wilcoxon test and unpaired Student's *t*-tests. The threshold for statistical significance was set at $P < 0.05$.

Results

Identification of DMGs, DEGs, and MeDEGs

In DNA methylation dataset, DMP and DMR analyses were performed. A total of 125 DMRswere identified at probe number greater than or equal to 10, containing 1059 cytosine-guanine (CpG) sites (Figure 2A). Subsequently, 10,617 DMGs were found, containing 2290 down-regulated genes and 8327 up-regulated genes (Figure 2B).

Altogether 2196 DEGs were acquired among the PDR and normal samples in the total PDR RNA-seq dataset, comprising 1006 up-regulated and 1190 down-regulated expressed genes in PDR samples (Figure 2C). Meanwhile, 27 DEGs, including 8 up-regulated and 19 down-regulated DEGs were identified between NPDR and normal groups in NPDR RNA-seq dataset (Figure 2D). Subsequently, 1283 PDR MeDEGs and 12 NPDR MeDEGs were identified in the intersection of DMGs with PDR DEGs and NPFR DEGs, respectively (Figure 2E).

PPI Network Analysis, MR Analysis, and Machine Learning-Based Integration Screened Biomarker of PDR

Based on the 1283 PDR MeDEGs, the PPI networks of the top 50 genes were constructed by 3 algorithms (MNC, Degree, and EPC), respectively (Supplementary Figure S1). Then, after the crossover of the first 50 genes of the 3 algorithms, 41 intersection genes were determined (Figure 3A).

After MR analysis, 4 intersection genes (CD44, PTPRC, MYD88, and TNF) were found to be apparently causally associated with DR and all of them were protective factors for PDR (odds ratio (OR) < 1, $P < 0.05$). In the scatter plot, the slopes of all 4 intersection genes were negative (Figure 3B and Supplementary Figure S2A–C). The overall effect value of the 4 intersection genes on DR was less than 0 in the forest plots (Figure 3C and Supplementary Figure S2D–F). The funnel plots showed that the SNPs loci of the 4 intersection genes were roughly left-right symmetric on both sides of the IVW line, in

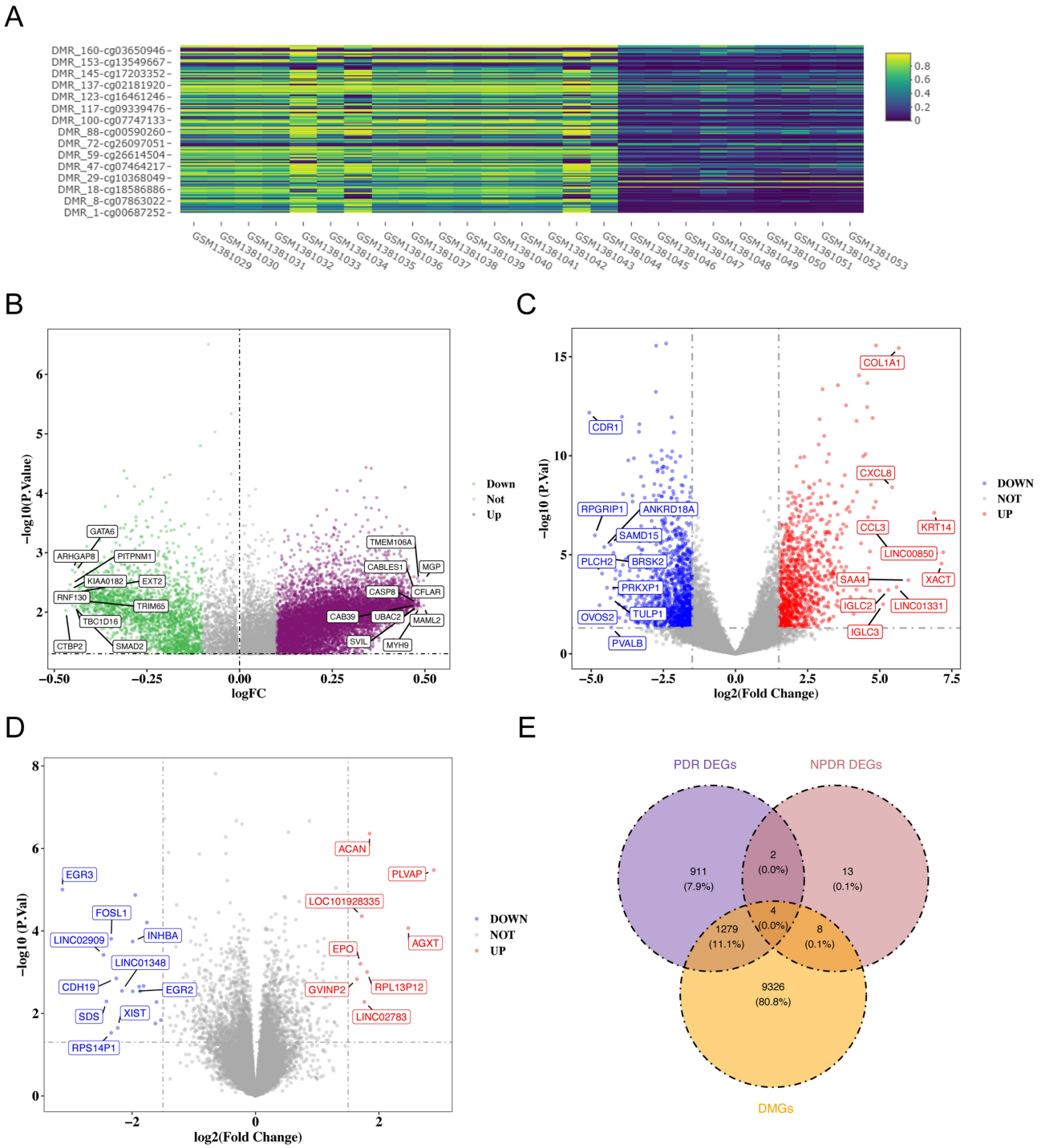

Figure 2 DMGs, DEGs, and MeDEGs in DR. (**A**) Heatmap for 1059CpGs in all DMRs. (**B**) The volcano plot of DMGs. The purple dots represent up-regulation and the green dots represent down-regulation. (**C**) The volcano plot of DEGs in the total PDR RNA-seq dataset. (**D**) The volcano plot of DMGs in the NPDR RNA-seq dataset. The red dots represent up-regulation and the blue dots represent down-regulation. (**E**) The intersection Venn diagram of DMGs, PDR DEGs and NPDR DEGs.

accordance with Mendel's second law (Figure 3D and Supplementary Figure S2G–I). Subsequently, in the heterogeneity test, the P values for CD44, PTPRC, and MYD88 were all greater than 0.05. These 3 genes were used for subsequent analysis. In the test for horizontal pleiotropy, the P values for MYD88, CD44, and PTPRC were all greater than 0.5, indicating no horizontal pleiotropy. As each SNP was progressively eliminated, and the remaining SNPs had less effect on the PDR (Figure 3E and Supplementary Figure S2J and K). In the MR Steiger test, the directions were all "TRUE" and the P-values

https://doi.org/10.2147/DMSO.S557119

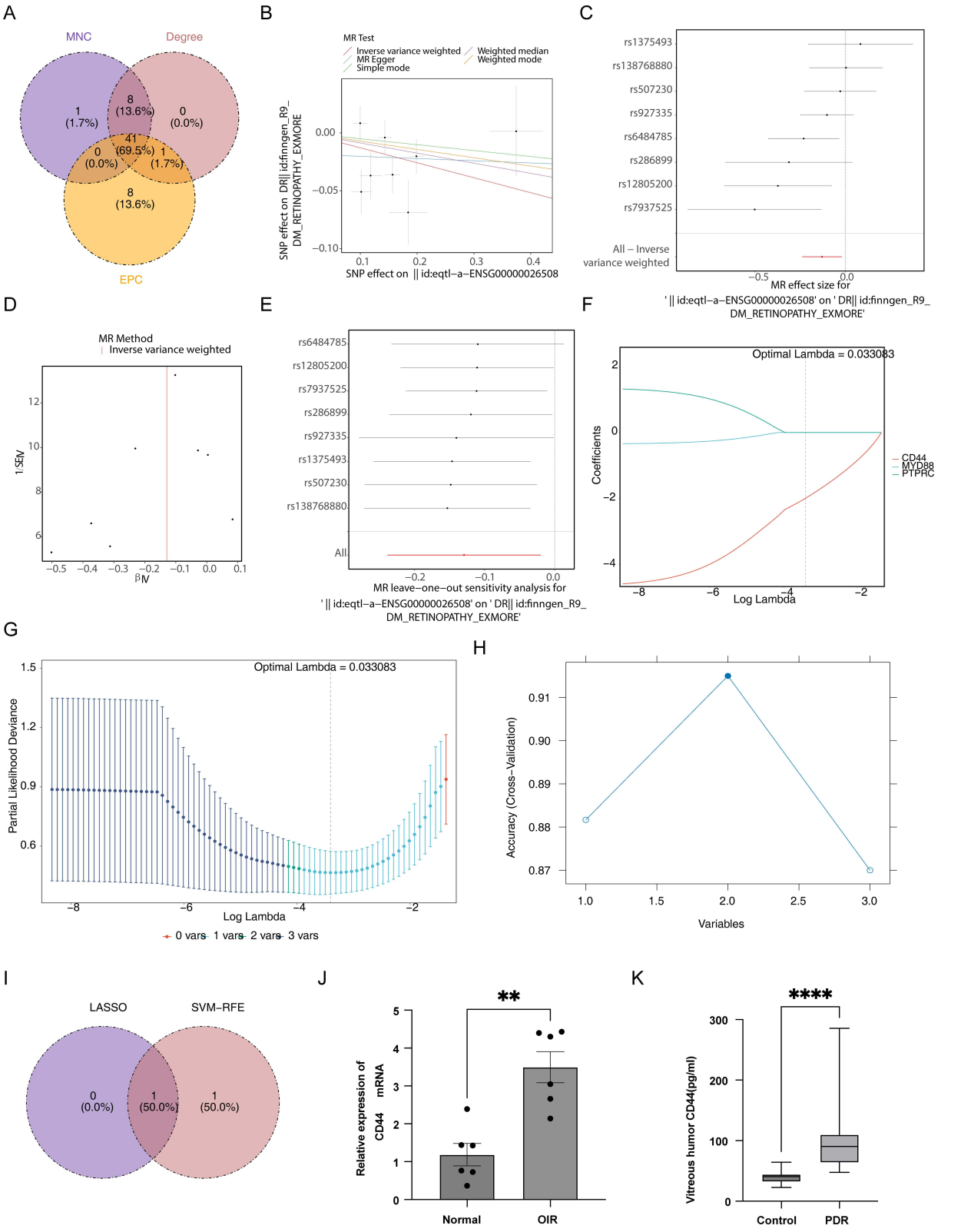

Figure 3 PPI network analysis, MR analysis, and machine learning-based integration screened biomarkers of PDR. (**A**) Venn diagram representing the intersection of the first 50 genes of the 3 algorithms. (**B**) Scatter plot showing the causal effect of CD44 on the risk of DR. (**C**) Forest plot showing the causal effect of each SNP on the risk of DR. (**D**) Funnel plots to visualize heterogeneity of MR estimates for the effect of CD44 on DR. (**E**) Forest plot showing the leave-one-out sensitivity analysis of the effect of CD44 on DR. (**F** and **G**) LASSO regression analysis was performed to select optimal feature genes. (**H**) The SVM-RFE analysis was used to screened optimal feature genes. (**I**) Venn diagram showing a biomarker shared by LASSO and SVM-REF algorithms. (**J**) The qRT-PCR analysis of CD44 mRNA levels in the retina of OIR model (n=6). (**K**) Comparison of vitreous humor CD44 levels between macular hole patients (Control group; n = 20) and PDR patients (PDR group; n = 19). Data are represented as box plots. Data represent the mean ± SEM. **P < 0.01, ****P < 0.0001, by independent samples *t*-test.

were all less than 0.05, indicating that the direction of causality between these 3 genes and PDR was correct. Finally, CD44, PTPRC, and MYD88 were identified as genes with significant causal and stable results with PDR.

Based on the 3 genes, in the LASSO model, 1 gene (CD44) was not penalized as 0 when lambda was set to a minimum value of 0.033083, thus CD44 was feature gene (Figure 3F and G). In the SVM-RFE analysis, 2 feature genes (CD44 and MYD88) were identified when the model error rate was lowest (Figure 3H). After the crossover of the feature genes obtained by the 2 algorithms, CD44 was obtained and considered as biomarker of PDR (Figure 3I). To verify the expression of CD44, we conducted qRT-PCR analysis on retinas collected from mouse OIR model. The results were consistent with the bioinformatics analysis, CD44 was highly expressed in the retinas of OIR model compared to the control (Figure 3J). Furthermore, the vitreous humor CD44 level was also higher in the PDR group (102.11 pg/mL, IQR 65.75–108.21 pg/mL) than in the Control group (39.29 pg/mL, IQR 32.08–44.19 pg/mL; Figure 3K).

MR Analysis and Machine Learning-Based Integration Screened Biomarkers of NPDR

Subsequently, the expression profiles of 12 NPDR MeDEGs were subjected to the machine learning-based integration to filter biomarkers of NPDR. The LASSO model with lambda.min equal to 0.01694229 included 8 genes whose coefficients were not penalized as 0, which were ACAN, PLVAP, EGR3, HBEGF, INHBA, CDH19, PCK1, and LAMC2 (Figure 4A and B). SVM-RFE identified 4 genes (ACAN, PLVAP, AGXT, and HBEGF) (Figure 4C). ACAN, PLVAP, and HBEGF were commonly identified by LASSO and SVM-RFE, thus they were biomarkers of NPDR (Figure 4D). To verify the expression of ACAN, PLVAP, and HBEGF, we conducted qRT-PCR analysis on retinas collected from T1DM and T2DM mouse. The results demonstrated ACAN mRNA level was up-regulated in the retinas of T1DM mice compared to the control (Figure 4E). PLVAP mRNA level was up-regulated in the retinas of T1DM and T2DM mice compared to the control (Figure 4F and G). HBEGF mRNA level was down-regulated in the retinas of T1DM mice compared to the control (Figure 4H).

The GSEA and ceRNAs Network of Biomarkers

GSEA showed that CD44 was significantly enriched for 96 pathways including ribosome, lysosome, antigen processing and presentation, and so on (Figure 5A and Supplementary Table 1). The 30 pathways were significantly enriched by ACAN, containing ribosome, and oxidative phosphorylation, among others (Figure 5B and Supplementary Table 2). The 40 pathways, such as ribosome, aminoacyl tRNA biosynthesis and axon guidance, were significantly enriched by PLVAP (Figure 5C and Supplementary Table 3). Furthermore, 28 pathways were obviously enriched by HBEGF, such as ribosome, P53 signaling pathway and cytokine-cytokine receptor interaction (Figure 5D and Supplementary Table 4). To further illuminate the potential interactions of DR lncRNAs, miRNAs and mRNAs, we constructed a lncRNA-miRNA-mRNA regulatory network based on the biomarkers. Top 5 miRNAs with high confidence targeting CD44, ACAN, HBEGF, and PLVAP were predicted, respectively. Based on 14 miRNAs, 23 lncRNAs were predicted (Figure 5E).

Distinct Cell-Specific Expression Patterns of Biomarkers via scRNA-Seq Analysis

Because of the difficulty in obtaining single-cell data from the retinas of patients with DR, we first conducted scRNA-seq analysis on GSE165784 to explore the cell-specific expression of biomarkers in fibrous membrane. The pre-processing procedure of scRNA-seq analysis, including QC, subsequent analysis, PCA, UMAP, is shown in Supplementary Figure S3. Upon cellular annotation, 6 cell types were identified, which were B cells, endothelial cells, macrophages, monocytes, T cells, and tissue stem cells (Figure 6A). Next, we investigated the cell-specific expression patterns of four biomarkers (Figure 6B). CD44 exhibited high expression in monocytes/macrophages. Moreover, ACAN demonstrated predominant expression in stem cells, while PLVAP exhibited marked expression in endothelial cells. HBEGF showed elevated expression levels in both monocytes and endothelial cells.

In addition, we also conducted scRNA-seq analysis on GSE205123 to explore the cell-specific expression of biomarkers in the retina of DM mice. The 6 cell types were identified, which were dendritic cells (DCs), granulocyte-monocyte progenitors (GMPs), hematopoietic stem cells (HSCs), macrophages, neuroepithelial cells, and neutrophils (Figure 6C). CD44 exhibited high expression in macrophages, while PLVAP and HBEGF were mainly enriched in neutrophils (Figure 6D).

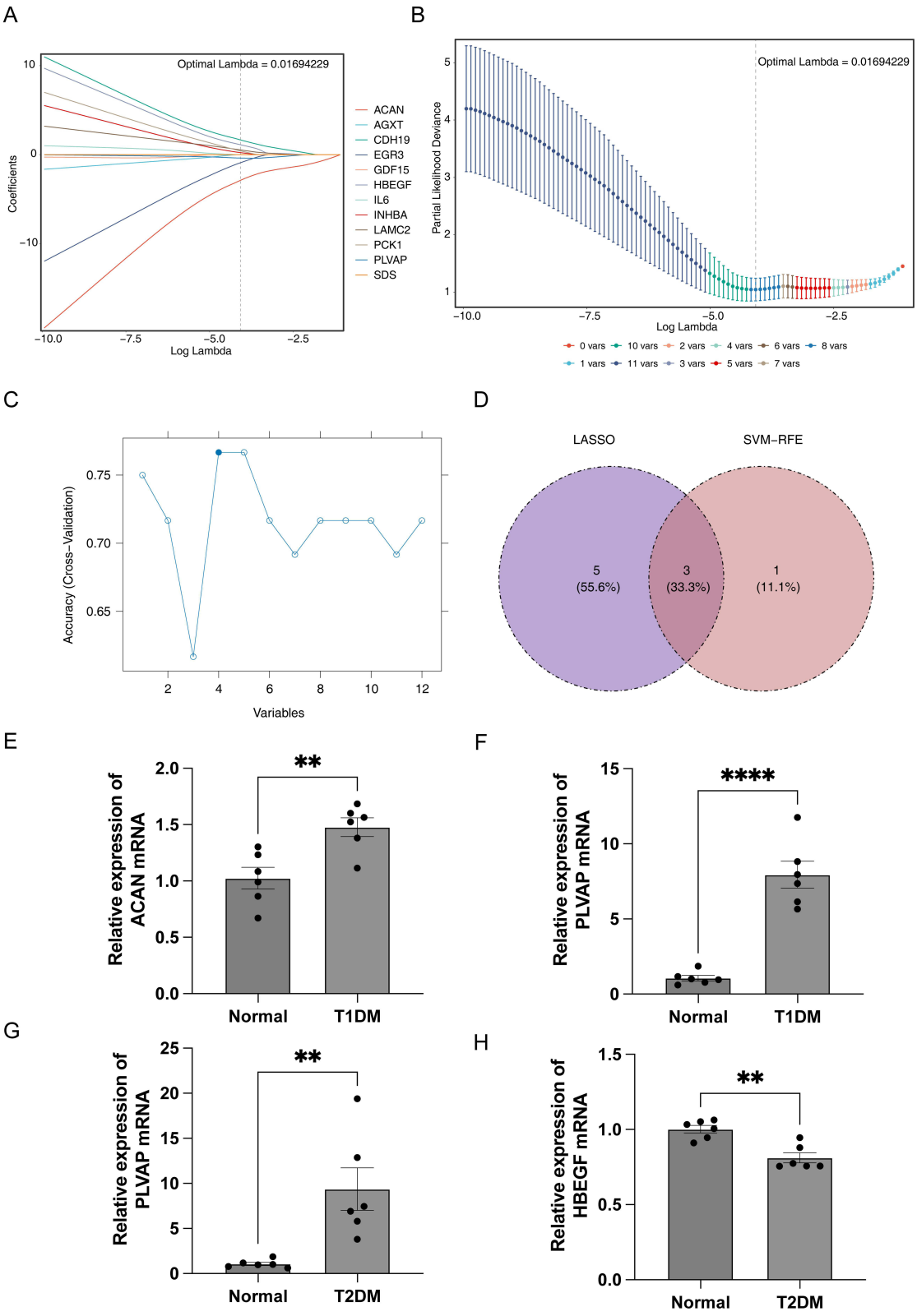

Figure 4 MR analysis and machine learning-based integration screened biomarkers of NPDR. (**A** and **B**) LASSO regression analysis was performed to select optimal feature genes. (**C**) The SVM-RFE analysis was used to screened optimal feature genes. (**D**) Venn diagram showing 3 biomarkers shared by LASSO and SVM-REF algorithms. (**E**) The qRT-PCR analysis of ACAN mRNA levels in the retina of T1DM mice (n=6). (**F**) The qRT-PCR analysis of PLVAP mRNA levels in the retina of T1DM mice (n=6). (**G**) The qRT-PCR analysis of PLVAP mRNA levels in the retina of T2DM mice (n=6). (**H**) The qRT-PCR analysis of HBEGF mRNA levels in the retina of T2DM mice (n=6). Data represent the mean ± SEM. **P < 0.01, ****P < 0.0001, by independent samples *t*-test.

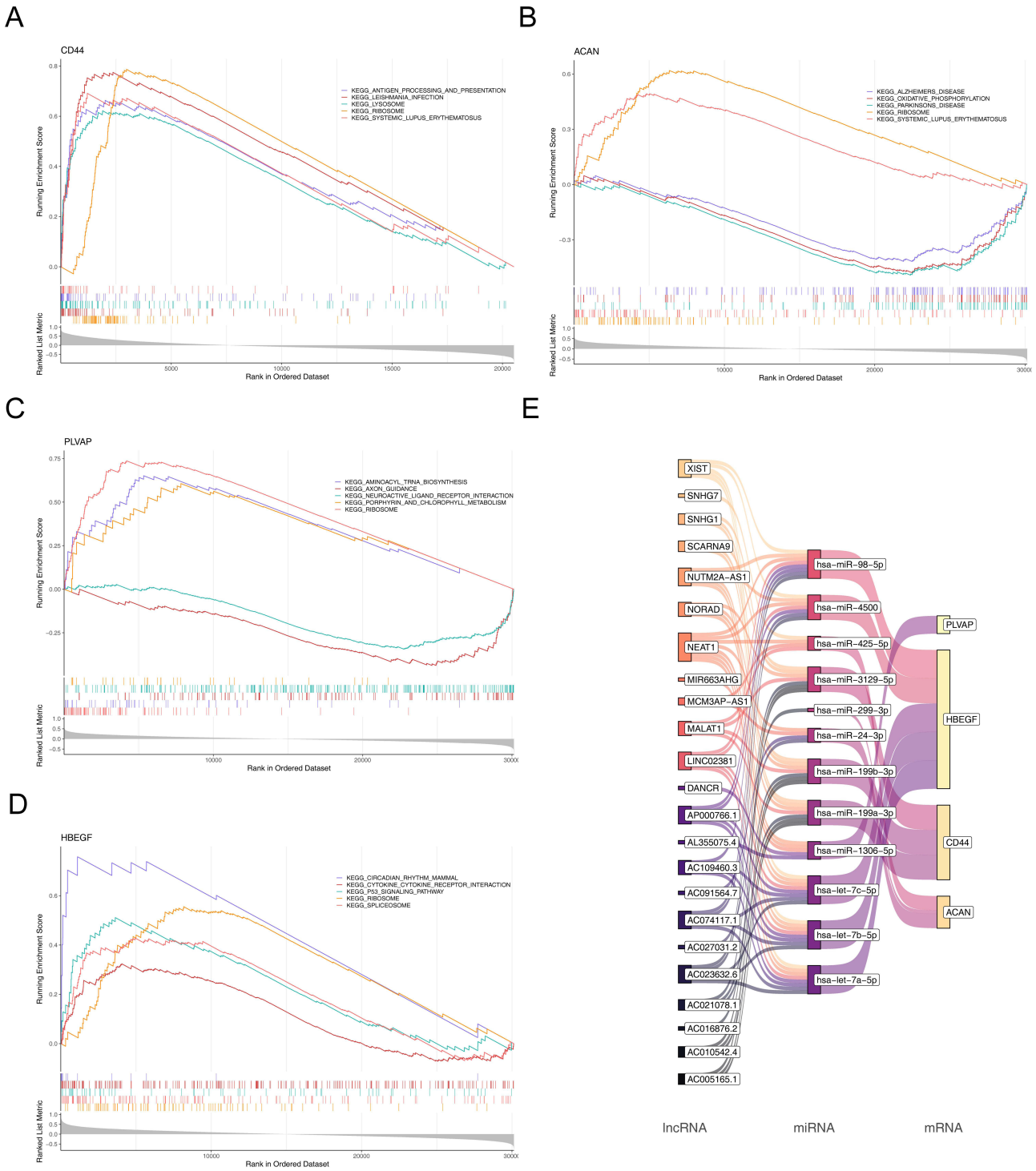

Figure 5 The GSEA enrichment analysis and ceRNAs network of biomarkers. (**A**–**D**) Top five signaling pathways that are significantly enriched in the high expression of CD44 (**A**), ACAN (**B**), PLVAP (**C**), and HBEGF (**D**). (**E**) lncRNA-miRNA-mRNA regulatory network based on the biomarkers.

Assessment of Immune Cell Infiltration

Then, we performed immune infiltration analysis to identify the immune cell types involved in the pathology of PDR and NPDR. Compared to those in the control group, the score of activated CD4 T cell, activated CD8 T cell, activated dendritic cell, CD56dim natural killer (NK) cells, central memory CD4 T cell, central memory CD8 T cell, effector memory CD8 T cell, gamma delta T cell, immature B cell, immature dendritic cell, macrophage, mast cell, myeloid-

https://doi.org/10.2147/DMSO.S557119

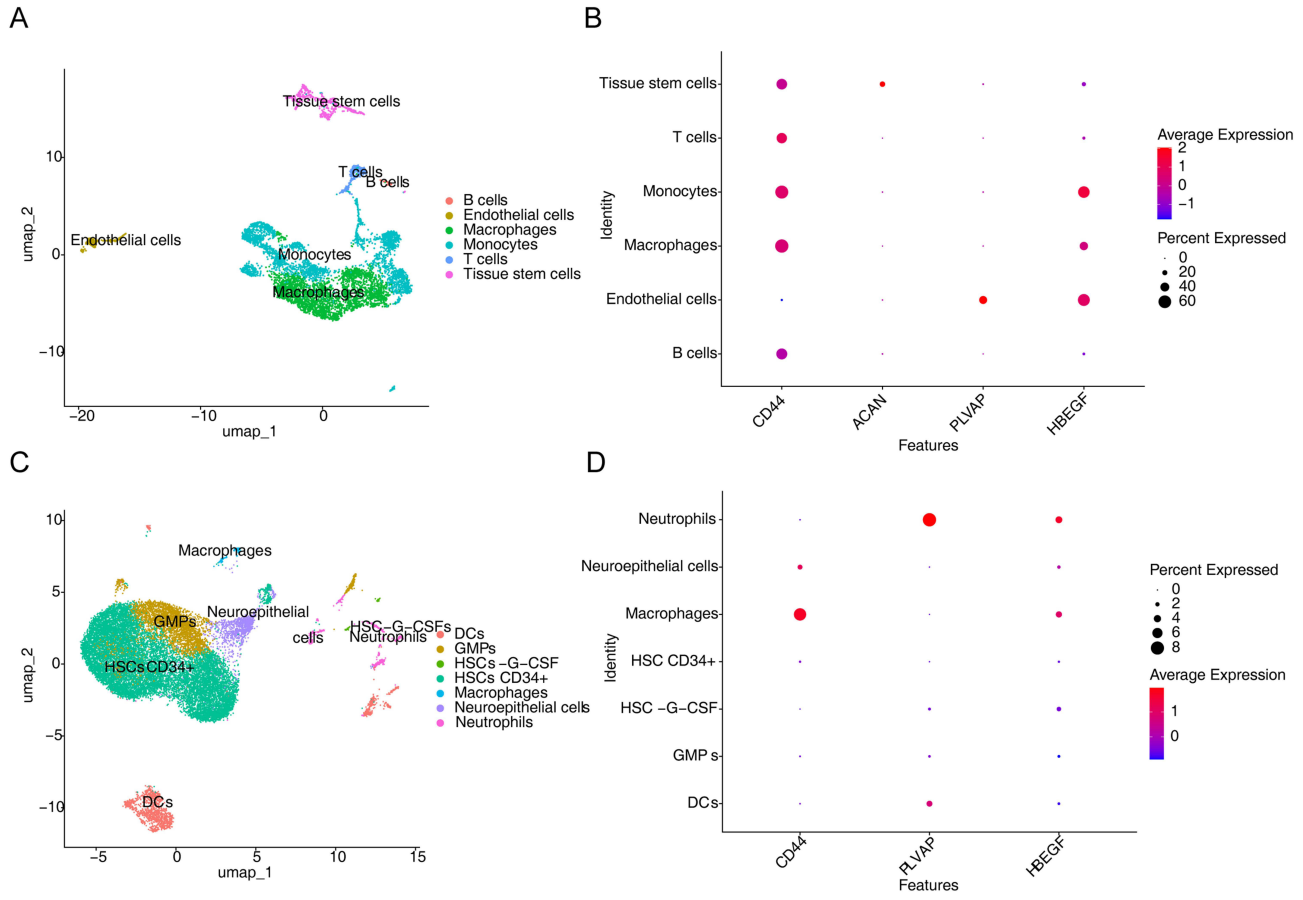

Figure 6 The scRNA-seq analysis of biomarkers. (**A**) Annotation on clusters using specific gene markers in fibrous membrane. (**B**) The bubble chart displays the expression of CD44, ACAN, PLVAP and HBEGF in cell clusters. (**C**) Annotation on clusters using specific gene markers in the retina of DM mice. (**D**) The bubble chart displays the expression of CD44, PLVAP and HBEGF in cell clusters.

derived suppressor cell (MDSC), memory B cell, monocyte, NK cell, NK T cell, plasmacytoid dendritic cell, regulatory T (Treg T) cell, T follicular helper cell, Type 1 T helper cell, Type 2 T helper cell were relatively high in the PDR samples ($P < 0.05$) (Figure 7A and B). Compared to those in the control group, the score of type 17 T helper (Th17) cell was relatively low, while the score of effector memory CD4 T cell was relatively high in the NPDR samples ($P < 0.05$) (Figure 7C and D).

In addition, correlations between differential immune cells were explored and found that most of the differential immune cells were positively correlated with each other in the total PDR dataset (cor > 0.3, $P < 0.05$) (Figure 7E). Among them, there was a maximum positive correlation between MDSCs and activated dendritic cells (cor = 0.94, $P < 0.05$) (Figure 7E). In the NPDR dataset, the correlation between type 17 T helper cell and effector memory CD4 T cell was not significant (cor = −0.21). CD44 was positively correlated with the remaining differential immune cells in the total PDR dataset (cor > 0.3, $P < 0.05$), except for type 2 T helper cells (Figure 7F). In the NPDR dataset, effector memory CD4 T cells had the largest positive correlation with ACAN (cor = 0.33, $P < 0.05$) and the largest negative correlation with HBEGF (cor = −0.48, $P < 0.05$) (Figure 7G).

Diagnostic Value of Biomarkers

We executed an ROC analysis to further evaluate the predictive power of biomarkers for PDR and NPDR. CD44 demonstrated an extremely robust diagnostic value (AUC = 0.964) in distinguishing PDR from control (Figure 8A). ACAN, PLVAP, and HBEGF also have potential for high differentiation between NPDR and control and the AUC values were 0.897, 0.770 and 0.724 (Figure 8B–D). In addition, the nomogram was constructed by ACAN, PLVAP, and HBEGF

to predict occurrence of NPDR (Figure 8E). In the calibration curve, the P-value was 0.553 in the Hosmer-Lemeshow (HL) test, and the predicted curve was closer to the ideal curve (Figure 8F). The AUC value of the nomogram was 0.91 (Figure 8G). Furthermore, the net benefit value of nomogram was greater than 0 and was higher than the net benefit values of the individual biomarkers. (Figure 8H). These findings suggested that nomogram had a high degree of prediction power for NPDR occurrence.

Biomarker-Drug Interaction

To explore the interaction between biomarkers and potential therapeutic drugs of DR, the biomarker-drug interaction network was visualized by Cytoscape. As shown in Figure 9 and Supplementary Table 5, a variety of FDA approved and unapproved drugs can affect the expression of CD44 and HBEGF.

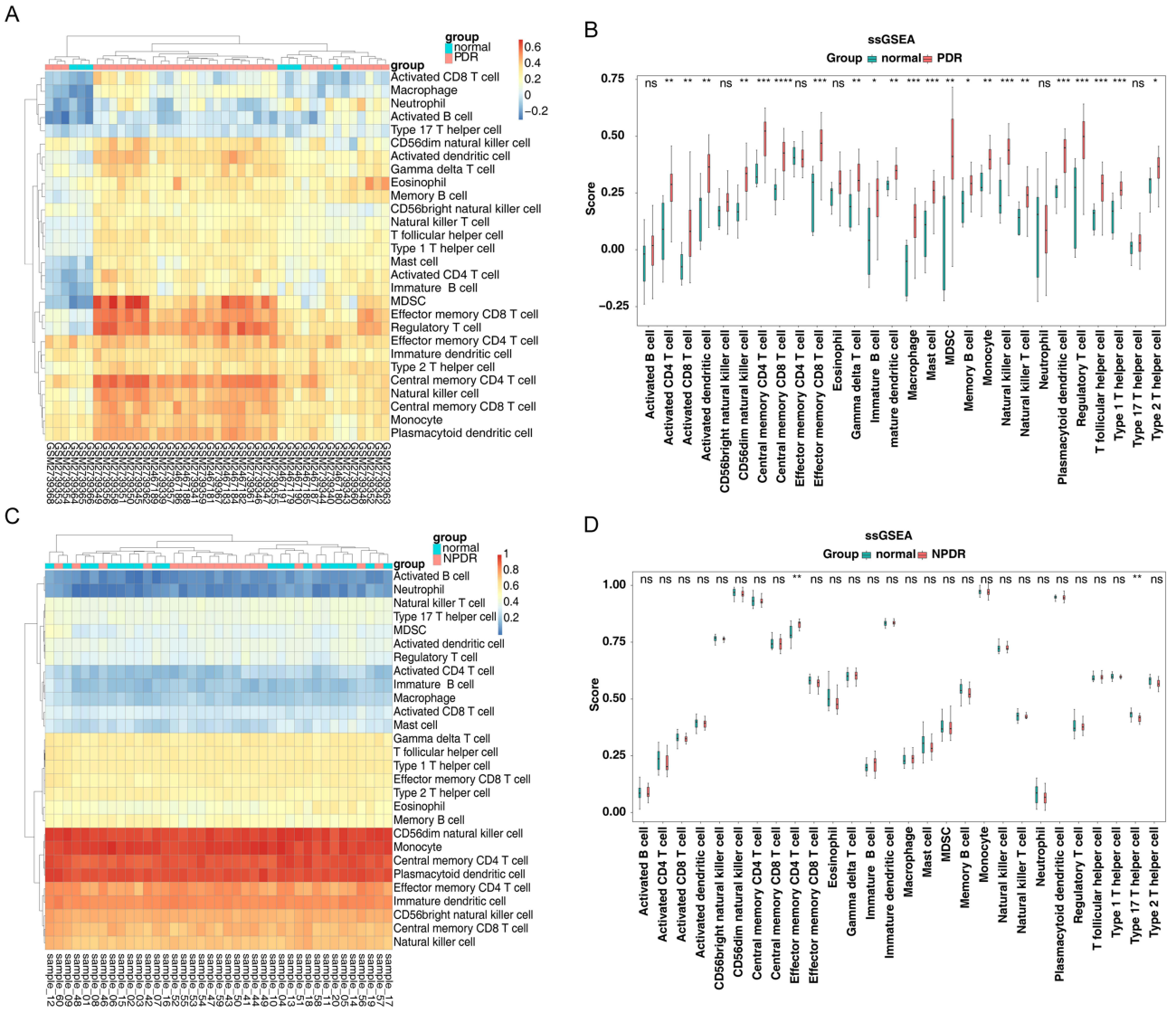

Figure 7 Continued.

https://doi.org/10.2147/DMSO.S557119

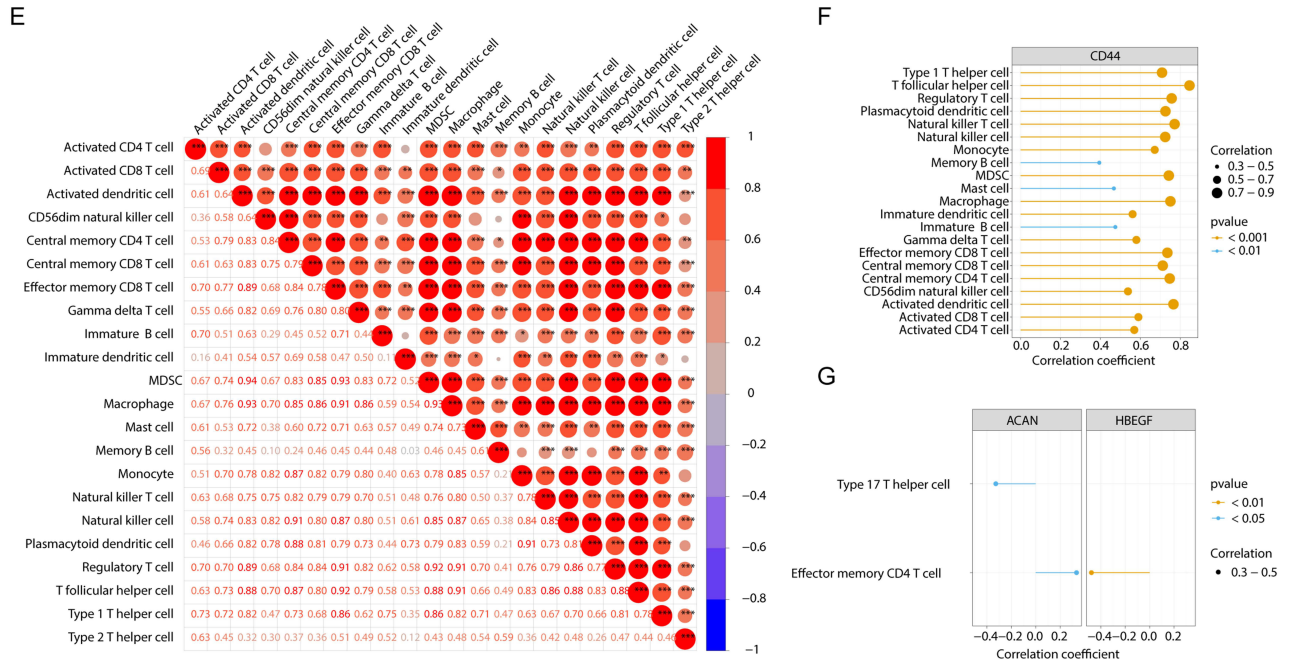

Figure 7 Immune cell infiltration analysis in DR patients. (**A**) Heatmap exhibited distribution of 28 types of immune cells between the PDR and control group. (**B**) The boxplot displaying the differences in the score of the 28 immune cells in PDR patients versus controls. (**C**) Heatmap exhibited distribution of 28 types of immune cells between the NPDR and control group. (**D**) The boxplot displaying the differences in the score of the 28 immune cells in NPDR patients versus controls. (**E**) Correlation matrix of all 28 types of immune cell subtype compositions. (**F**) Correlation between CD44 and differential immune cells. (**G**) Correlation between ACAN, HBEGF and differential immune cells. ns, no significance, *P< 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

Discussion

DR is one of the most serious neurovascular complications of DM, its occurrence involves a complex biological process, including augmentation of the polyol pathway, protein kinase C activation, increased oxidative stress, formation of advanced glycation end product, and upregulation of growth factor and adhesion molecules.[14] Therefore, it is crucial to identify the biomarkers and signaling pathways involved in the pathogenetic process in DR. Multi-omics approaches, combined with bioinformatics analysis, offer a novel perspective for uncovering the molecular mechanism underlying DR. DNA methylation plays a critical role in DR progression.[8] Because DNA methylation precedes pathological changes,[15] aberrant DNA methylation may serve as a potential biomarker for the early diagnosis of DR.[16] Hui et aldetermined that abnormal methylation of the NLRP3 inflammation-related gene promoter could elevate the risk of DR.[17] In addition, DNA methylation status also have potential as a biomarker to predict the progression of DR.[18] Studies have shown that abnormal methylation of tumor necrosis factor, chitinase 3-like protein 1, chimerin 2 and gastric inhibitory polypeptide receptor may serve as prospective biomarkers of PDR.[8,19,20] DNA methylation is also implicated in the complex pathological processes of DR, including xidative stress, inflammation, neovascularization and metabolic memory.[17,19,21–23] Thus, targeting DNA methylation could emerge a novel therapeutic strategy for DR. However, research on the methylation of DR-related genes remains limited. Identifying MeDEGs in DR through multi-omics approaches holds great significance for advancing the understanding of its mechanisms, diagnosis and treatment.

In this study, we identified 10,617 DMGs in DNA methylation dataset, 2196 DEGs in the total PDR dataset and 27 DEGs in NPDR dataset. After the intersection of the two types of data, 1283 PDR MeDEGs and 12 NPDR MeDEGs were found. Then, we screened 4 biomarkers between DR and the control group via bioinformatics methods. Our results showed CD44 was identified as biomarkers of PDR, while ACAN, PLVAP, HBEGF were identified as biomarkers for NPDR via MR analysis and machine learning-based integration. CD44, as a widely expressed cell surface adhesion molecule, promotes the occurrence and development of pathological angiogenesis by regulating the proliferation, migration, adhesion, invasion and communication with the microenvironment of endothelial cell.[24–28] Meanwhile, CD44 and VEGF exhibit a synergistic effect in promoting pathological neovascularization.[29] The protein encoded by

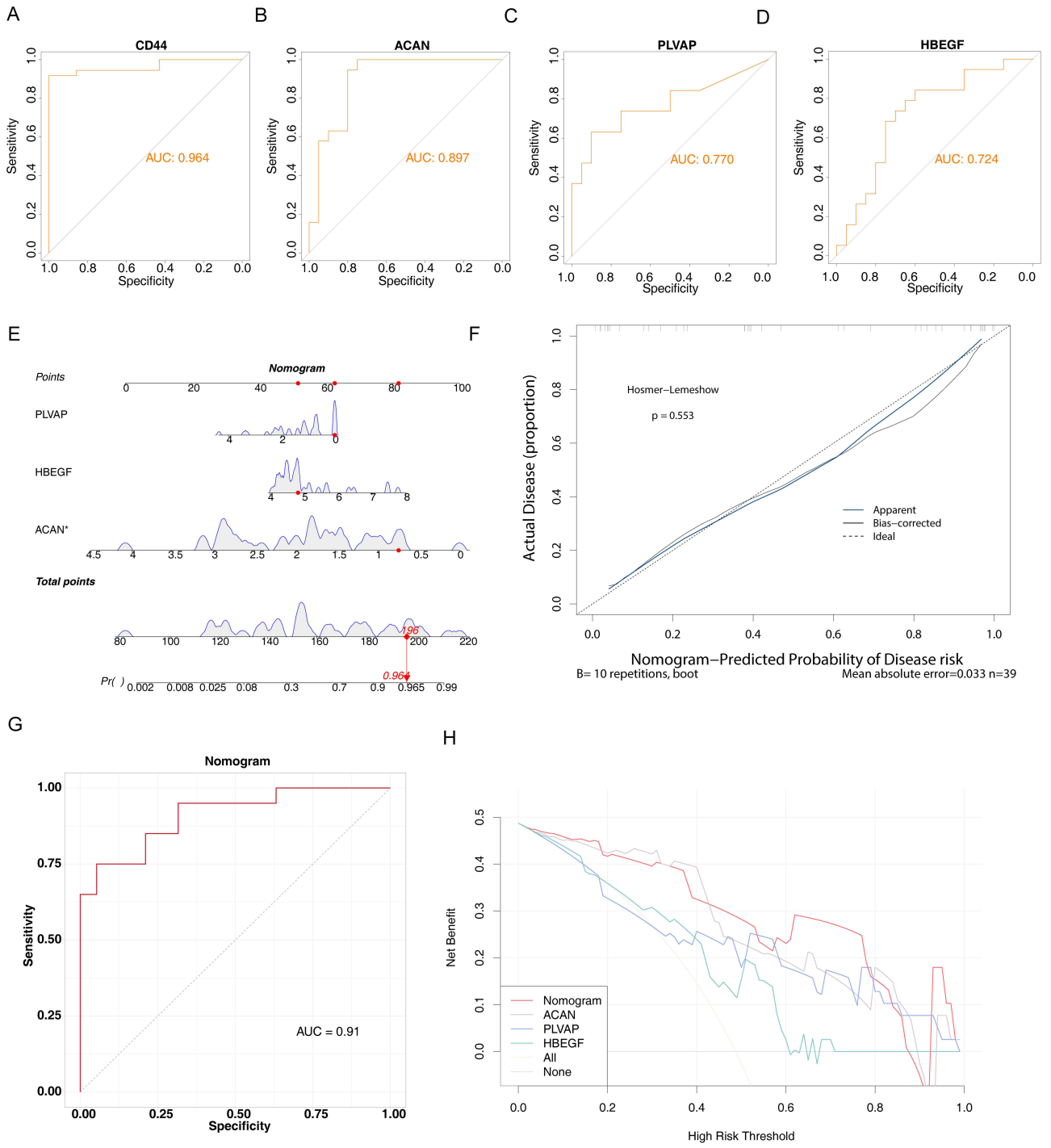

Figure 8 Diagnostic value of biomarkers. (**A**–**D**) The ROC curve was applied to verify accurateness of biomarkers. (**E**) Diagnostic nomogram of three biomarkers (ACAN, PLVAP, and HBEGF). (**F**) The calibration curves of nomogram. (**G**) The ROC curve was applied to verify accurateness of nomogram. (**H**) Decision curve was applied to verify net benefit value of nomogram.

ACAN is an important component of extracellular matrix (ECM) and contribute to the regulation of vascular ECM through diverse domain interactions.[30] Its accumulation may drive fibrotic tissue formation, vascular basement membrane thickening, and inflammation in DR.[31,32] PLVAP, an endothelial cell-specific protein, is expressed exclusively under pathological conditions associated with impaired barrier function.[33] Additionally, PLVAP acts as a downstream effector of VEGF and is implicated in increased vascular permeability and blood-retinal barrier (BRB) breakdown.[33,34]

https://doi.org/10.2147/DMSO.S557119

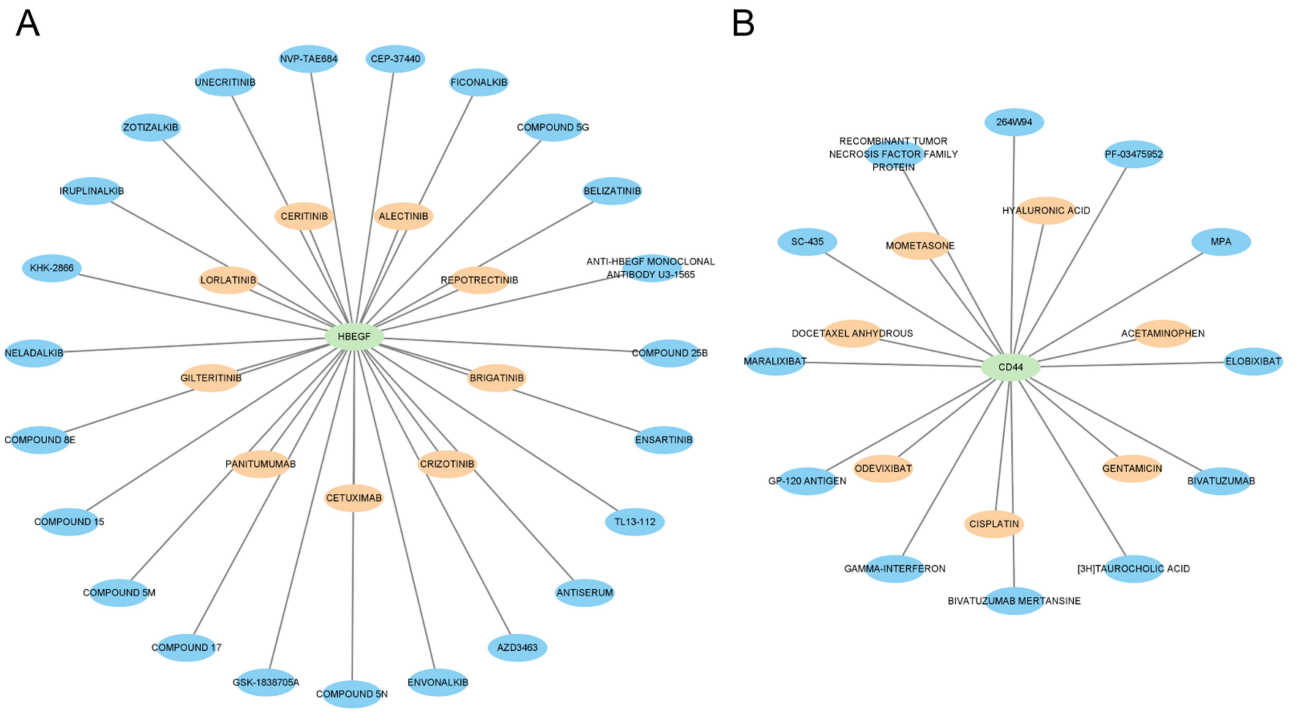

Figure 9 Interaction network between biomarkers and potential drugs. (**A**) Potential drugs targeting HBEGF. (**B**) Potential drugs targeting CD44. Green nodes represent the biomarkers (HBEGF and CD44); Orange nodes represent the FDA approved drugs; blue nodes represent the FDA unapproved drugs.

HBEGF is an epidermal growth factor receptor ligand, which plays an important role in angiogenesis, wound healing, oxidative stress, cell migration and epithelial-mesenchymal transition.[35,36] Leveraging scRNA-seq analysis, we uncovered the cell-specific expression patterns of biomarkers. In the retina of STZ-induced type 1 diabetic mice, PLVAP and HBEGF were mainly enriched in neutrophils. In the retinal fibroproliferative membranes of PDR patients, CD44 were mainly enriched in monocytes/macrophages. Neutrophils, which are frontline cells of the innate immune system, are typically associated with the initial phases of inflammation.[37,38] On the one hand, increased neutrophil infiltration and excessive neutrophil-endothelium adhesion lead to retinal inflammation and vascular leukostasis, contributing to the progression of DR.[39] On the other hand, neutrophils, through the release of neutrophil extracellular traps (NETs), target pathological angiogenesis for clearance and promote the regeneration of functional vessels.[38] However, uncontrolled retinal neutrophils and abundant NETs may potentially exacerbate the disruption of the blood-retinal barrier.[40,41] As the primary cellular component of the mononuclear phagocyte system, macrophages play a critical role in the pathogenesis and progression of DR.[42,43] During early-stage DR, macrophages polarize into pro-inflammatory M1 phenotypes, which drive retinal inflammation through the secretion of pro-inflammatory cytokines.[44] In later stages, M2-polarized macrophages shift to promoting pathological neovascularization and fibrovascular membrane formation via VEGF secretion.[45] Furthermore, senescent macrophages exhibit characteristic M2 surface markers while concurrently displaying a pro-inflammatory secretory profile. This paradoxical phenotype may explain the persistence of intraocular inflammation in PDR, despite the presence of fibrosis and neovascularization.[46] Gene therapy targeting neutrophils and macrophages may provide a new therapy strategy for DR.

The results of GSEA revealed that these biomarkers were associated with the ribosome, lysosome, oxidative phosphorylation, aminoacyl tRNA biosynthesis, and the P53 signaling pathway. Aberrantly expressed lncRNAs and miRNAs play pivotal roles in modulating target molecules and pathogenetic pathways in DR.[47–49] Based on biomarkers, we constructed a ceRNA network and identified specific lncRNAs and miRNAs.

Accumulating evidence has indicated that the activation of immune cells in the retina contribute to the pathogenesis of DR.[50] Leukocyte adhesion to the microvasculature results in endothelial cell loss and BRB breakdown.[51] Microglia, a class of mononuclear phagocytes, is involved in retinal neovascularization and inflammatory processes.[42,46,52] T lymphocytes appear

to negatively regulate retinal neovascularization.[53,54] Th17 cell infiltration and Treg cell imbalance were observed in the retina of DR patients and mouse models,[55,56] and these cells were involved in inflammation and abnormal angiogenesis.[57,58] Elevated levels of B-cell-produced antibodies were also observed in the vitreous humor of DR patients.[59] These results were generally consistent with our findings. The score of most of immune cells was relatively high in PDR samples. The score of type 17 T helper (Th17) cell was relatively low, while the score of effector memory CD4 T cell was relatively high in the NPDR samples. Moreover, the correlation analysis between biomarkers and immune cells revealed that CD44 was positively correlated with the differential immune cells except for type 2 T helper cells. Effector memory CD4 T cells exhibited the strongest positive correlation with ACAN and the strongest negative correlation with HBEGF. These findings further highlight the critical role of immune cell infiltration in the pathogenesis of DR.

Finally, we also evaluated the diagnostic performance of biomarkers. The ROC curve showed that biomarkers exhibited excellent diagnostic efficacy. Subsequently, a diagnostic nomogram model was established based on biomarkers associated with NPDR. ROC curve, decision curve and calibration curves collectively validated the accuracy and sensitivity of the model in distinguishing between NPDR and normal groups.

However, our research has several limitations. Firstly, we conducted cross-species scRNA-seq analysis on biomarkers of NPDR because of data sparsity. Secondly, the retinal tissue specificity of the selected IV remains unverified in MR analysis. Thirdly, due to the lack of human retinal samples, the expression of biomarkers in the retina was validated only in mouse models. Fourthly, due to the lack of external validation, diagnostic models have the risk of overfitting. Future research should include large sample size with detailed clinical information to validate biomarkers and diagnostic models, along with molecular biology studies to further explore the pathogenesis of DR.

Conclusions

In conclusion, we reported potential methylation-regulated biomarkers (CD44, ACAN, PLVAP, and HBEGF) involved in in the pathophysiology of DR and elucidated their molecular functions by integrated bioinformatics analysis. These findings provide valuable insights for further investigation into the mechanisms underlying the DR development.

Abbreviations

AUC, area under the curve; ceRNAs, Competing endogenous RNAs; DMGs, differentially methylated genes; DMP, differential methylation probe; DMR, differential methylation region; DR, diabetic retinopathy; ELISA, Enzyme-linked immunosorbent assay; eQTL, expression quantitative trait locus; GSEA, gene set enrichment analysis; MeDEGs, methylation-regulated differentially expressed genes; MR, mendelian randomization; NPDR, non-proliferative diabetic retinopathy; PCA, principal component analysis; PDR, proliferative diabetic retinopathy; PPI, protein-protein interaction; QC, quality control; qRT-PCR, Quantitative real-time polymerase chain reaction; ROC, receiver operating curve; scRNA-seq, Single-cell RNA sequencing; SNPs, single nucleotide polymorphisms.

Data Sharing Statement

The data that supports the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval and Consent to Participate

The experimental protocol was established, according to the ethical guidelines of the Helsinki Declaration and the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research and was approved by the Human Ethics Committee of Tianjin Medical University Eye Hospital and Tianjin Medical University Eye Institute Laboratory Animal Center (TJYY20231202894). Written informed consent was obtained from individual or guardian participants. Authors have adhered to the ARRIVE guidelines.

Author Contributions

Qingbo Li: Writing - Original Draft, Data Curation, Investigation. Xu Zhou: Writing - Original Draft, Data Curation, Validation. Ying Wang: Writing - Original Draft, Data Curation, Resources. Yi Peng: Writing - Original Draft, Validation, Project Administration. Juping Liu: Methodology, Supervision, Writing - Review & Editing. Xiaorong Li:

Conceptualization, Methodology, Supervision, Writing - Review & Editing. Yan Shao (corresponding author): Conceptualization, Methodology, Supervision, Resources, Writing - Review & Editing. All authors gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by grants from the project of central government supports for Tibet development (XZ202301YD0029C), and National Natural Science Foundation of China (82360209).

Disclosure

The authors declare that they have no competing interests.

References

1. Sun H, Saeedi P, Karuranga S, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabet Res Clin Pract*. 2022;183:109119. doi:10.1016/j.diabres.2021.109119
2. Steinmetz JD, Bourne RRA, Briant PS. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the right to sight: an analysis for the global burden of disease study. *Lancet Glob Health*. 2021;9:e144–e60. doi:10.1016/s2214-109x(20)30489-7
3. Tan TE, Wong TY. Diabetic retinopathy: looking forward to 2030. *Front Endocrinol*. 2022;13:1077669. doi:10.3389/fendo.2022.1077669
4. Duh EJ, Sun JK, Stitt AW. Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. *JCI Insight*. 2017;2. doi:10.1172/jci.insight.93751
5. Couturier A, Rey PA, Erginay A, et al. Widefield OCT-angiography and fluorescein angiography assessments of nonperfusion in diabetic retinopathy and edema treated with anti-vascular endothelial growth factor. *Ophthalmology*. 2019;126:1685–1694. doi:10.1016/j.ophtha.2019.06.022
6. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38:23–38. doi:10.1038/npp.2012.112
7. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13:484–492. doi:10.1038/nrg3230
8. Cai C, Meng C, He S, et al. DNA methylation in diabetic retinopathy: pathogenetic role and potential therapeutic targets. *Cell Biosci*. 2022;12:186. doi:10.1186/s13578-022-00927-y
9. Lu J, Huang Y, Zhang X, Xu Y, Nie S. Noncoding RNAs involved in DNA methylation and histone methylation, and acetylation in diabetic vascular complications. *Pharmacol Res*. 2021;170:105520. doi:10.1016/j.phrs.2021.105520
10. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol*. 2015;16:37. doi:10.1186/s13059-015-0600-x
11. Su M, Pan T, Chen QZ, et al. Data analysis guidelines for single-cell RNA-seq in biomedical studies and clinical applications. *Mil Med Res*. 2022;9:68. doi:10.1186/s40779-022-00434-8
12. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18:83. doi:10.1186/s13059-017-1215-1
13. Karczewski KJ, Snyder MP. Integrative omics for health and disease. *Nat Rev Genet*. 2018;19:299–310. doi:10.1038/nrg.2018.4
14. Whitehead M, Wickremasinghe S, Osborne A, Van Wijngaarden P, Martin KR. Diabetic retinopathy: a complex pathophysiology requiring novel therapeutic strategies. *Expert Opin Biol Ther*. 2018;18:1257–1270. doi:10.1080/14712598.2018.1545836
15. Zhang X, Zhao L, Hambly B, Bao S, Wang K. Diabetic retinopathy: reversibility of epigenetic modifications and new therapeutic targets. *Cell Biosci*. 2017;7:42. doi:10.1186/s13578-017-0167-1
16. Bansal A, Pinney SE. DNA methylation and its role in the pathogenesis of diabetes. *Pediatr Diabetes*. 2017;18:167–177. doi:10.1111/pedi.12521
17. Chen H, Zhang X, Liao N, et al. Identification of NLRP3 inflammation-related gene promoter hypomethylation in diabetic retinopathy. *Invest Ophthalmol Vis Sci*. 2020;61:12. doi:10.1167/iovs.61.13.12
18. Duraisamy AJ, Radhakrishnan R, Seyoum B, Abrams GW, Kowluru RA. Epigenetic modifications in peripheral blood as potential noninvasive biomarker of diabetic retinopathy. *Transl Vis Sci Technol*. 2019;8:43. doi:10.1167/tvst.8.6.43
19. Agardh E, Lundstig A, Perfilyev A, 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;13:182. doi:10.1186/s12916-015-0421-5
20. Zhao T, Su Z, Li Y, Zhang X, You Q. Chitinase-3 like-protein-1 function and its role in diseases. *Signal Transduct Target Ther*. 2020;5:201. doi:10.1038/s41392-020-00303-7
21. Berdasco M, Gómez A, Rubio MJ, et al. DNA methylomes reveal biological networks involved in human eye development, functions and associated disorders. *Sci Rep*. 2017;7:11762. doi:10.1038/s41598-017-12084-1
22. Sundrani DP, Reddy US, Joshi AA, et al. Differential placental methylation and expression of VEGF, FLT-1 and KDR genes in human term and preterm preeclampsia. *Clin Clin Epigenet*. 2013;5:6. doi:10.1186/1868-7083-5-6
23. Mishra M, Kowluru RA. The role of DNA methylation in the metabolic memory phenomenon associated with the continued progression of diabetic retinopathy. *Invest Ophthalmol Vis Sci*. 2016;57:5748–5757. doi:10.1167/iovs.16-19759
24. Chen L, Fu C, Zhang Q, He C, Zhang F, Wei Q. The role of CD44 in pathological angiogenesis. *FASEB j*. 2020;34:13125–13139. doi:10.1096/fj.202000380RR
25. Yu Q, Stamenkovic I. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. *Genes Dev*. 2000;14:163–176. doi:10.1101/gad.14.2.163
26. Matou-Nasri S, Gaffney J, Kumar S, Slevin M. Oligosaccharides of hyaluronan induce angiogenesis through distinct CD44 and RHAMM-mediated signalling pathways involving Cdc2 and gamma-adducin. *Int J Oncol*. 2009;35:761–773. doi:10.3892/ijo_00000389

27. Park D, Kim Y, Kim H, et al. Hyaluronic acid promotes angiogenesis by inducing RHAMM-TGFβ receptor interaction via CD44-PKCδ. *Mol Cells*. 2012;33:563–574. doi:10.1007/s10059-012-2294-1
28. Zhang SS, Hu JQ, Liu XH, et al. Role of moesin phosphorylation in retinal pericyte migration and detachment induced by advanced glycation endproducts. *Front Endocrinol*. 2020;11:603450. doi:10.3389/fendo.2020.603450
29. Shen WY, Yu MJ, Barry CJ, Constable IJ, Rakoczy PE. Expression of cell adhesion molecules and vascular endothelial growth factor in experimental choroidal neovascularisation in the rat. *Br J Ophthalmol*. 1998;82:1063–1071. doi:10.1136/bjo.82.9.1063
30. Koch CD, Lee CM, Apte SS. Aggrecan in cardiovascular development and disease. *J Histochem Cytochem*. 2020;68:777–795. doi:10.1369/0022155420952902
31. Roy S, Bae E, Amin S, Kim D. Extracellular matrix, gap junctions, and retinal vascular homeostasis in diabetic retinopathy. *Exp Eye Res*. 2015;133:58–68. doi:10.1016/j.exer.2014.08.011
32. Roy S, Amin S, Roy S. Retinal fibrosis in diabetic retinopathy. *Exp Eye Res*. 2016;142:71–75. doi:10.1016/j.exer.2015.04.004
33. Bosma EK, van Noorden CJF, Schlingemann RO, Klaassen I. The role of plasmalemma vesicle-associated protein in pathological breakdown of blood-brain and blood-retinal barriers: potential novel therapeutic target for cerebral edema and diabetic macular edema. *Fluids Barriers CNS*. 2018;15:24. doi:10.1186/s12987-018-0109-2
34. Wisniewska-Kruk J, van der Wijk AE, van Veen HA, et al. Plasmalemma vesicle-associated protein has a key role in blood-retinal barrier loss. *Am J Pathol*. 2016;186:1044–1054. doi:10.1016/j.ajpath.2015.11.019
35. Hult EM, Gurczynski SJ, O'Dwyer DN, et al. Myeloid- and epithelial-derived heparin-binding epidermal growth factor-like growth factor promotes pulmonary fibrosis. *Am J Respir Cell Mol Biol*. 2022;67:641–653. doi:10.1165/rcmb.2022-0174OC
36. Kim S, Subramanian V, Abdel-Latif A, Lee S. Role of heparin-binding epidermal growth factor-like growth factor in oxidative stress-associated metabolic diseases. *Metab Syndr Relat Disord*. 2020;18:186–196. doi:10.1089/met.2019.0120
37. Metzemaekers M, Gouwy M, Proost P. Neutrophil chemoattractant receptors in health and disease: double-edged swords. *Cell Mol Immunol*. 2020;17:433–450. doi:10.1038/s41423-020-0412-0
38. Binet F, Cagnone G, Crespo-Garcia S, et al. Neutrophil extracellular traps target senescent vasculature for tissue remodeling in retinopathy. *Science*. 2020:369. doi:10.1126/science.aay5356
39. Barouch FC, Miyamoto K, Allport JR, et al. Integrin-mediated neutrophil adhesion and retinal leukostasis in diabetes. *Invest Ophthalmol Vis Sci*. 2000;41:1153–1158.
40. Kessenbrock K, Krumbholz M, Schönermarck U, et al. Netting neutrophils in autoimmune small-vessel vasculitis. *Nat Med*. 2009;15:623–625. doi:10.1038/nm.1959
41. Li T, Qian Y, Li H, et al. Cellular communication network factor 1 promotes retinal leakage in diabetic retinopathy via inducing neutrophil stasis and neutrophil extracellular traps extrusion. *Cell Commun Signal*. 2024;22:275. doi:10.1186/s12964-024-01653-3
42. Hu A, Schmidt MHH, Heinig N. Microglia in retinal angiogenesis and diabetic retinopathy. *Angiogenesis*. 2024;27:311–331. doi:10.1007/s10456-024-09911-1
43. Kinuthia UM, Wolf A, Langmann T. Microglia and inflammatory responses in diabetic retinopathy. *Front Immunol*. 2020;11:564077. doi:10.3389/fimmu.2020.564077
44. Yao Y, Li J, Zhou Y, et al. Macrophage/microglia polarization for the treatment of diabetic retinopathy. *Front Endocrinol*. 2023;14:1276225. doi:10.3389/fendo.2023.1276225
45. Wang X, Xu C, Bian C, et al. M2 microglia-derived exosomes promote vascular remodeling in diabetic retinopathy. *J Nanobiotechnology*. 2024;22:56. doi:10.1186/s12951-024-02330-w
46. Li Q, Wang P, Gong Y, et al. α-Klotho prevents diabetic retinopathy by reversing the senescence of macrophages. *Cell Commun Signal*. 2024;22:449. doi:10.1186/s12964-024-01838-w
47. Wu Y, Jia K, Wu H, et al. A comprehensive competitive endogenous RNA network pinpoints key molecules in diabetic retinopathy. *Mol Med Rep*. 2019;19:851–860. doi:10.3892/mmr.2018.9715
48. Li J, Li C, Zhao Y, et al. Integrated bioinformatics analysis for novel miRNAs markers and ceRNA network in diabetic retinopathy. *Front Genet*. 2022;13:874885. doi:10.3389/fgene.2022.874885
49. Wang T, Cheng MY, Shan MY, et al. Construction of a competitive endogenous RNA network related to exosomes in diabetic retinopathy. *Comb Chem High Throughput Screen*. 2023;26:576–588. doi:10.2174/1386207325666220610122114
50. Adamis AP, Berman AJ. Immunological mechanisms in the pathogenesis of diabetic retinopathy. *Semin Immunopathol*. 2008;30:65–84. doi:10.1007/s00281-008-0111-x
51. Sun WJ, An XD, Zhang YH, et al. The ideal treatment timing for diabetic retinopathy: the molecular pathological mechanisms underlying early-stage diabetic retinopathy are a matter of concern. *Front Endocrinol*. 2023;14:1270145. doi:10.3389/fendo.2023.1270145
52. Yu Y, Chen H, Su SB. Neuroinflammatory responses in diabetic retinopathy. *J Neuroinflammation*. 2015;12:141. doi:10.1186/s12974-015-0368-7
53. Ishida S, Usui T, Yamashiro K, et al. VEGF164-mediated inflammation is required for pathological, but not physiological, ischemia-induced retinal neovascularization. *J Exp Med*. 2003;198:483–489. doi:10.1084/jem.20022027
54. Ishida S, Yamashiro K, Usui T, et al. Leukocytes mediate retinal vascular remodeling during development and vaso-obliteration in disease. *Nat Med*. 2003;9:781–788. doi:10.1038/nm877
55. Sigurdardottir S, Zapadka TE, Lindstrom SI, et al. Diabetes-mediated IL-17A enhances retinal inflammation, oxidative stress, and vascular permeability. *Cell Immunol*. 2019;341:103921. doi:10.1016/j.cellimm.2019.04.009
56. Yang TT, Song SJ, Xue HB, Shi DF, Liu CM, Liu H. Regulatory T cells in the pathogenesis of type 2 diabetes mellitus retinopathy by miR-155. *Eur Rev Med Pharmacol Sci*. 2015;19:2010–2015.
57. Deliyanti D, Talia DM, Zhu T, et al. Foxp3(+) tregs are recruited to the retina to repair pathological angiogenesis. *Nat Commun*. 2017;8:748. doi:10.1038/s41467-017-00751-w
58. Han JM, Patterson SJ, Speck M, Ehses JA, Levings MK. Insulin inhibits IL-10-mediated regulatory T cell function: implications for obesity. *J Immunol*. 2014;192:623–629. doi:10.4049/jimmunol.1302181
59. Liu B, Hu Y, Wu Q, et al. Qualitative and quantitative analysis of B-cell-produced antibodies in vitreous humor of type 2 diabetic patients with diabetic retinopathy. *J Diabetes Res*. 2020;2020:4631290. doi:10.1155/2020/4631290