

# Biomarkers of Arginine Methylation in Diabetic Nephropathy: Novel Insights from Bioinformatics Analysis

Yiming Guan, Xiayan Yin, Liyan Wang, Zongli Diao, Hongdong Huang\*, Xueqi Wang<sup>id</sup>\*

Department of Nephrology, Beijing Friendship Hospital, Capital Medical University, Beijing, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Xueqi Wang, Hongdong Huang, Department of Nephrology, Beijing Friendship Hospital, Capital Medical University, 95 Yong An Road, Xi Cheng District, Beijing, 100050, People's Republic of China, Email Xueqi wang32@hotmail.com; Huanghd1126@126.com

**Background:** Diabetic nephropathy (DN) is a severe complication of diabetes influenced by arginine methylation. This study aimed to elucidate the role of protein arginine methylation-related genes (PRMT-RGs) in DN and identify potential biomarkers.

**Methods:** Differentially expressed genes in two GEO datasets (GSE30122 and GSE104954) were integrated with 9 PRMT-RGs. Candidate genes were identified using WGCNA and differential expression analysis, then screened using support vector machine-recursive feature elimination and least absolute shrinkage and selection operator. Biomarkers were defined as genes with consistent differential expression across both datasets. Regulatory networks were constructed using the miRNet and Network Analyst databases. Gene set enrichment analysis was performed to identify the signaling pathways in which the biomarkers were enriched in DN. Different immune cells in DN were identified using immune infiltration analysis. Meanwhile, drug prediction and molecular docking identified potential DN therapies. Finally, qRT-PCR and immunohistochemistry validated two biomarkers in STZ-induced DN mice and DN patients.

**Results:** Two biomarkers (FAM98A and FAM13B) of DN were identified in this study. The molecular regulatory network revealed that FAM98A and FAM13B were co-regulated by 6 microRNAs and 1 transcription factor and were enriched in signaling pathways. Immune infiltration and correlation analyses revealed that FAM98A and FAM13B were involved in developing DN along with PRMT-RGs and immune cells. The expression levels of Fam98a and Fam13b were significantly upregulated in the kidneys of DN mice revealed by qRT-PCR analysis. The expression levels of FAM98A were significantly upregulated in the kidneys of DN patients revealed by immunohistochemistry staining. Molecular docking showed that estradiol and rotenone exerted potential therapeutic effects on DN by targeting FAM98A.

**Conclusion:** Comprehensive bioinformatics analysis revealed that FAM98A and FAM13B were potential DN biomarkers correlated with PRMT-RGs and immune cells. This study provided useful insights for elucidating the molecular mechanisms and developing targeted therapy for DN.

**Keywords:** diabetic nephropathy, protein arginine methylation-related genes, biomarkers, bioinformatics analysis

## Introduction

The 10-year cumulative mortality rate of diabetic nephropathy (DN), which affects approximately one-third of patients with diabetes mellitus, is as high as 31.1%.<sup>1</sup> The pathogenic mechanisms of DN are diverse and complex, encompassing hemodynamic alterations (such as imbalanced arteriolar resistance), metabolic factors leading to oxidative stress, disruptions in cellular signal transduction, transcription factors (TFs), and pro-inflammatory factors.<sup>2</sup> These pathways induce apoptosis and podocyte loss and decrease the glomerular filtration rate (GFR), contributing to the onset and progression of DN. Clinically, DN is diagnosed based on renal biopsy, an invasive procedure. Currently, the major therapeutics for DN are renin-angiotensin-aldosterone system inhibitors.<sup>3</sup> Various randomized clinical trials have demonstrated the cardiorenal benefits of drugs, such as sodium-glucose

cotransporter 2 inhibitors or mineralocorticoid receptor antagonist in patients with DN.<sup>4–8</sup> However, the incidence of DN is increasing. This indicates that only pharmaceutical interventions may not yield favorable outcomes in DN. Although DN has been extensively studied, the underlying pathological mechanisms have not been elucidated. Further studies are needed to identify the biomarkers and elucidate the potential molecular mechanisms of DN, which will aid in the development of novel diagnostic and therapeutic strategies for DN.

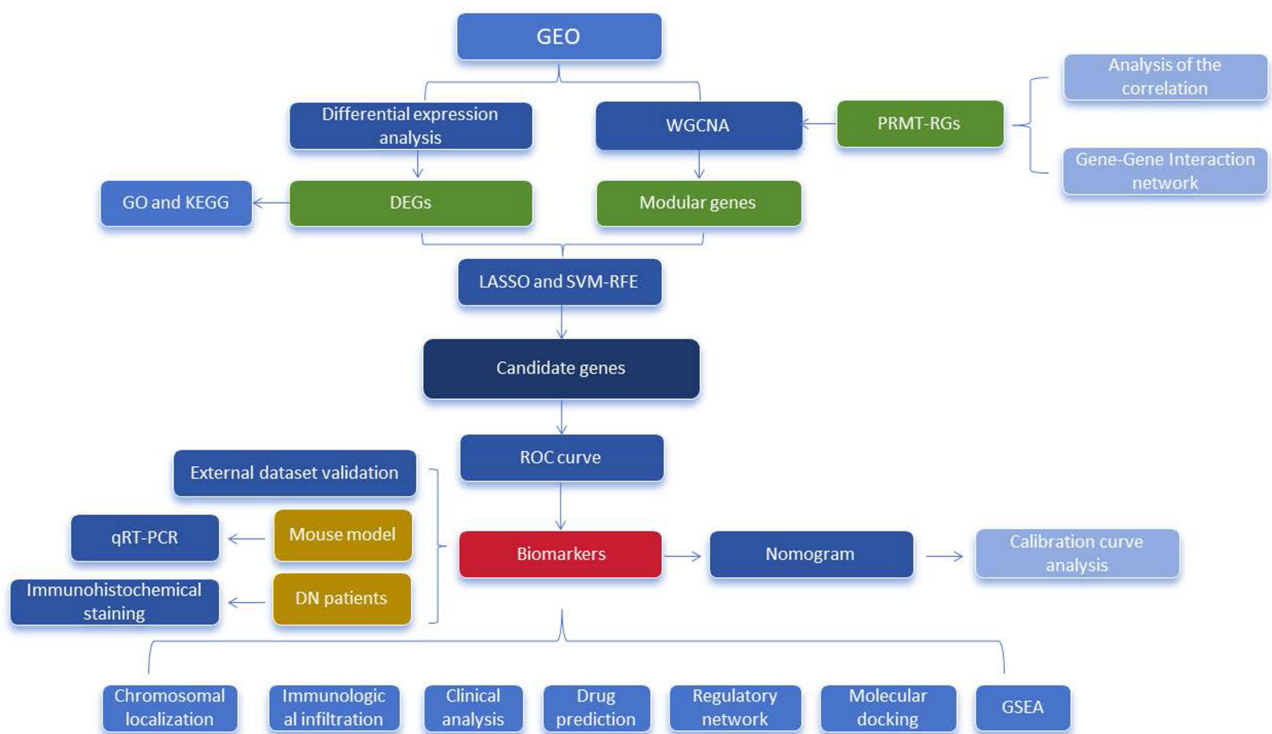
Arginine, a basic  $\alpha$ -amino acid residue, constitutes various proteins, including histones. Protein arginine methylation (PRMT) regulates several biological processes, including transcription, splicing, RNA biology, DNA damage response, and cell metabolism,<sup>9</sup> which are dysregulated in kidney diseases. Several studies have indicated the role of PRMTs in kidney diseases, including renal fibrosis, acute kidney injury (AKI), and DN. PRMT1, which mainly promotes asymmetric arginine methylation of histone and nonhistone proteins, may mediate renal fibroblast activation and renal fibrosis development.<sup>10</sup> In ischemia/reperfusion-induced AKI, the accumulation of asymmetric dimethylarginine, a metabolic product of PRMTs, activates oxidative stress, promoting tubular necrosis.<sup>11</sup> The upregulation of *PRMT1* levels in DN induces oxidative stress, apoptosis in mesangial cells, and diabetic kidney injury, whereas the knockdown of PRMT1 alleviates diabetic kidney injury.<sup>12,13</sup> These findings indicate the critical role of PRMTs in the development and maintenance of DN.

Bioinformatics is characterized by large-scale, high-dimensional, and imbalanced data sets. Machine learning techniques have proven highly effective in data preprocessing and feature extraction, significantly improving data quality.<sup>14</sup> These techniques have become crucial in various domains, including genomics, protein structure prediction, biological network analysis, drug discovery, and disease prediction and diagnosis, driving significant advances in the biological sciences.<sup>15–18</sup> Hongdong Han et al employed advanced machine learning algorithms such as LASSO, random forest (RF), and support vector machine recursive feature elimination (SVM-RFE) to accurately identify PRKAR2B and TGFBI as critical biomarkers for diabetic nephropathy (DN).<sup>19</sup> This work provides robust scientific evidence for the early diagnosis of DN and establishes a strong foundation for future research aimed at identifying therapeutic targets using bioinformatics and machine learning. It also paves the way for developing specific therapeutic strategies for DN, highlighting the substantial potential of bioinformatics and machine learning in precision medicine. This study focuses on utilizing advanced bioinformatics approaches to deeply investigate the molecular mechanisms of protein arginine methyltransferase-related genes (PRMT-RGs) in the pathogenesis of diabetic nephropathy (DN). We mined the GEO database to identify differentially expressed genes between DN and normal control groups and performed cross-comparisons with PRMT-RGs to pinpoint key candidate genes. Subsequently, we employed various machine learning algorithms to identify potential biomarkers and validated their expression patterns in a mouse DN model and DN patients. Furthermore, we utilized bioinformatics tools such as GO annotation, KEGG pathway analysis, protein-protein interaction network construction, and immune infiltration analysis to comprehensively elucidate the roles of these genes in the pathological processes of DN. Additionally, through potential drug prediction, this study aims to pave new avenues for DN treatment, thereby advancing the application and development of precision medicine in the DN field.

## Materials and Methods

### Data Sources

The GSE30122 and GSE104954 microarray expression datasets retrieved from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) served as the training and validation datasets, respectively. The GSE30122 dataset comprised the data of renal tubule tissues from 10 patients with DN and 12 healthy subjects on the GPL570 platform.<sup>20</sup> Meanwhile, the GSE104954 dataset comprised the data of renal tubule tissues from 7 patients with DN and 18 healthy subjects on the GPL22945 platform.<sup>21</sup> The following nine PRMT-RGs were obtained from the published literature: *PRMT1*, *PRMT2*, *PRMT3*, *CARM1* (also called *PRMT4*), *PRMT5*, *PRMT6*, *PRMT7*, *PRMT8*, and *PRMT9*.<sup>22</sup> The flow chart of this study is shown in Figure 1.



**Figure 1** The flow chart of this study.

## Analysis of the Correlation Between PRMT-RGs

To explore the interactions between PRMT-RGs, the gene-gene interaction network (GGI) of PRMT-RGs was constructed using the GeneMANIA network (<http://www.genemania.org>). The TOP 20 correlated genes were selected as the key nodes for displaying, and the TOP 7 significant pathways were presented. Furthermore, a heatmap was generated to illustrate the correlation between PRMT-RGs as determined using Spearman correlation coefficient.

## Weighted Gene Co-Expression Network Analysis (WGCNA)

The PRMT-RG scores were calculated for each sample in the GSE30122 dataset using single-sample gene set enrichment analysis (ssGSEA). WGCNA was performed with PRMT-RG score as the trait data using the WGCNA package (version 1.72–1).<sup>23</sup> The hierarchical clustering analysis was implemented, and outlier samples were removed. Subsequently, the scale-free topology fit index of 0.8 was used to determine the most suitable soft threshold power. The power adjacent function of Pearson's correlation matrix was used to transform filtered gene expression data into a topological overlap matrix (TOM). The TOM was transformed into a dissimilarity matrix from which a systematic clustering tree between genes was obtained. A dynamic shear tree was constructed with a minModuleSize of 100. To identify the modules associated with PRMT-RG scores, modules and PRMT-RG scores were subjected to Pearson's correlation analysis. The correlation between modules and PRMT-RG scores was displayed as a heatmap. The module with the most significant correlation with PRMT-RG scores was selected as the key module.

## DEG Identification

The “limma” (version 3.52.4)<sup>24</sup> package was used to screen the DEGs between the DN and control groups in the training dataset based on the following criteria: adjusted  $p < 0.05$ ;  $|\log_2\text{fold change} (\log_2\text{FC})| > 1$ . The DEGs were visualized using volcano plots and heatmaps using the packages “ggplot2” (version 3.4.1)<sup>25</sup> and “pheatmap” (version 2.14.0), respectively.<sup>26</sup>

## Enrichment Analysis of DEGs

DEGs were subjected to GO and KEGG enrichment analyses using the “clusterProfiler” (version 4.7.1) package<sup>27</sup> to examine the biological functions and signaling pathways associated with the DEGs (adj  $p < 0.05$ ). The results were visualized using the “GPlot” (version 1.0.2) package ( $p < 0.05$ ).<sup>28</sup>

## Identification of Candidate Genes

The DEGs were intersected with the key module genes of the key module using Venn diagram analysis. Next, feature genes were screened using the least absolute shrinkage and selection operator (LASSO) algorithm with the “Glmnet” (version 4.1.7) package<sup>29</sup> and support vector machine-recursive feature elimination (SVM-RFE) algorithm with the “e1071” package (version 1.7–12).<sup>30</sup> The overlapping genes in the two algorithms were identified as candidate genes.

## Diagnostic Performance Evaluation and Validation of Candidate Genes

The receiver operating characteristic (ROC) curve, which is the most central index for medical diagnostic tests, enables the evaluation of predictive model performance. The diagnostic efficacy of the candidate genes in the training and validation datasets was evaluated using the “pROC” package (version 1.18.0).<sup>31</sup> Intergroup difference analysis was performed to determine the expression patterns of the candidate genes in both training and validation datasets. Finally, the candidate genes that passed validation were considered biomarkers for DN in this study.

## Construction and Evaluation of Nomogram

The contribution of biomarkers to DN pathogenesis was presented using a nomogram. In the training dataset, the construction of nomograms was performed using the “rms” package (version 6.7–0).<sup>32</sup> Furthermore, the effectiveness of the nomogram was assessed using decision curve analysis (DCA).

## Chromosomal Localization and Clinical Analysis of Biomarkers

To clarify the location of biomarkers on human chromosomes, the “RCircos” package (version 1.2.2)<sup>33</sup> was used to visualize the distribution of biomarkers on chromosomes. The correlation between biomarker expression and clinical characteristics (GFR) was examined using the Nephroseq V5 database (<http://v5.nephroseq.org>).

## Gene Set Enrichment Analysis (GSEA)

The “psych” package (version 2.2.9)<sup>34</sup> was used to perform Spearman correlation analysis between the biomarkers and the genes in the training dataset. Subsequently, the genes were sorted according to the degree of correlation and subjected to GSEA using the “clusterProfiler” (version 4.7.1) package.<sup>27</sup> The reference gene sets were “c2.cp.reactome.v7.0.symbols.gmt” and “c5.go.v7.4.symbols.gmt” from Molecular Signatures Database (<https://www.gsea-msigdb.org/gsea/msigdb/>). The KEGG pathways in which the biomarkers were enriched were analyzed (adj  $p < 0.05$ ). The top 10 most significantly enriched pathways for biomarkers were displayed using the “enrichplot” (version 1.16.2) package.<sup>35</sup>

## Analysis of Immune Cell Infiltration in DN

ssGSEA, which determines the enrichment scores for 28 immune cells, was used to determine the differential immune infiltration status between control and DN samples.<sup>36</sup> Samples were filtered based on  $p < 0.05$ . A heatmap illustrating immunological infiltration in the samples was generated using the “corrplot” (version 0.92) package.<sup>37</sup> Further, the correlation of biomarkers with PRMT-RGs and different immune cells was presented using heatmaps.<sup>38</sup>

## Regulatory Network Construction and Targeted Drug Prediction

The miRNet web database (<https://www.mirnet.ca>) was used to identify potential biomarkers targeted by miRNAs. The miRNA-biomarker regulatory networks were visualized using Cytoscape. NetworkAnalyst online tool (<https://www.networkanalyst.ca>) was used to generate a gene regulation network for the TF-biomarker pair. Data on target TFs and genes were extracted from the ENCODE chromatin immunoprecipitation-sequencing data.



## Drug Prediction and Molecular Docking

Small-molecule drugs were identified using the biomarkers available at the Comparative Toxicogenomics Database (CTD) (<http://ctdbase.org/>). The networks depicting the interaction between biomarkers and drugs were visualized using Cytoscape (version 3.9.1). Docking of biomarkers and small-molecule drugs was performed to identify potential biomarker-targeting drugs. The three-dimensional structures of drugs and proteins were obtained from the PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and UniProt (<https://www.uniprot.org/>) databases, respectively. AutoDock Tools were applied to perform protein hydrogenation, charge calculation, and charge equilibrium. Docking of biomarkers and potential drugs was performed using AutoDock vina (version 1.5.7). The structure exhibiting the lowest binding free energy in the resulting output was selected. Finally, PyMol software (version 2.0) was used to visualize the docking results.<sup>39</sup>

## Construction of the Mouse Model

C57BL/6 male and female mice were obtained from Weitong Lihua Limited Company (Beijing, China) and housed in standardized conditions with a 12-h light/dark cycle. The animals were allowed to acclimatize for one week and randomly divided into the following two groups: experimental group, intraperitoneally administered with streptozotocin (STZ) (Sigma, CAS No. 18883–66-4, USA) at a dose of 120 mg/kg bodyweight; control group, administered with an equivalent volume of citrate buffer (Solarbio, China). The animals were allowed to fast for 12 h before drug administration. The blood glucose levels of >11.1 mmol/L on day 3 post-injection suggested successful modeling. Additionally, the blood glucose levels were measured on day 6 post-treatment to confirm stable hyperglycemia.

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

Total mRNA was extracted from five DN samples and five control samples using TRIzol (Invitrogen, Thermofisher, USA), following the manufacturer's instructions. The RNA samples were subjected to standard agarose gel electrophoresis. The quantity and quality of RNA samples were determined using a NanoPhotometer N50 (Implen, Germany). RNA was reverse-transcribed into complementary DNA (cDNA) using the SureScript-First-strand-cDNA-synthesis-kit (Servicebio, China), following the manufacturer's instructions. qRT-PCR analysis was performed using the 2× Universal Blue SYBR Green qPCR Master Mix (Servicebio, China) and the CFX Connect Real-Time Quantitative Fluorescence PCR Instrument (BIO-RAD, USA). The primer sequences used in qRT-PCR analysis are listed in Table 1. *Gapdh* was used as an endogenous control. The relative expression levels were determined using the  $2^{-\Delta\Delta CT}$  method.<sup>40</sup>

## Clinical Data Acquisition

All DN patients included in this study were admitted to the Department of Nephrology at Beijing Friendship Hospital and were diagnosed with DN through renal biopsy. The renal biopsies were conducted with each patient's informed consent. As a control group, age-matched tissues from patients with non-diabetic minimal change disease (MCD) were simultaneously collected. This study received approval from the Ethical Committee of Beijing Friendship Hospital (Ethics approval number: BFH20240528005/BFHHZS20240146).

**Table 1** Details of RT-qPCR Primers

Primers	Sequencing
FAM13B F	AGTGACCACGGGGATAGTGA
FAM13B R	CAGGAAACGGCTCTTCCCAT
FAM98A F	GGCAACTAACAGTCCAAGCG
FAM98A R	TTGAAAGACCTCACTACCGCC
M-GAPDH F	CCTTCCGTGTTCTACCCC
M-GAPDH R	GCCCAAGATGCCCTTCAGT

## Immunohistochemistry (IHC) Staining

Immunohistochemistry staining was carried out on 4  $\mu$ m thick paraffin-embedded kidney tissue sections. After deparaffinization, the fixed sections were microwaved for 20 minutes in sodium citrate buffer (pH 6.0) (Solarbio, China), followed by a 30-minute incubation with normal goat serum at room temperature. The sections were then treated with primary antibodies and left to incubate overnight at 4°C. The next day, the sections were brought back to room temperature and incubated with a secondary antibody (PV-9000, Zhongshan Goldenbridge, China) for 1 hour. Staining was then performed using the diaminobenzidine (DAB) reagent. Nuclei were counterstained with hematoxylin and washed to achieve a blue hue. Finally, the sections were mounted with neutral balsam. Images were captured using an Olympus CKX41 optical microscope (Japan). The primary antibodies used include Anti-FAM98A (PA5-90721, Invitrogen, USA).

## Statistical Analysis

Statistical analysis was performed using R programming language (version 4.2.2). Categorical data were compared using the Wilcoxon test. Differences were considered significant at  $p < 0.05$  unless specified otherwise.

## Results

### Expression Levels of PRMT1, PRMT2, PRMT3, CARM1, PRMT5, PRMT7 and PRMT8 Were Correlated in the Training Set

The GGI network revealed that nine PRMT-RGs interacted with various genes, including *HEMK1*, METTL family-encoding genes, and *ATP5CKMT*, which co-regulated the methyltransferase activity (Figure 2A). Analysis of the correlation between PRMT-RGs demonstrated a significant positive correlation between the *PRMT1* and *PRMT2* expression levels ( $R = 0.643$ ,  $p = 0.001$ ) and a significant negative correlation between the *PRMT8* and *PRMT3* expression levels ( $R = -0.635$ ,  $p = 0.001$ ) (Figure 2B). As *PRMT6* and *PRMT9* were not included in the training set, the remaining seven PRMT-RGs (*PRMT1*, *PRMT2*, *PRMT3*, *CARM1*, *PRMT5*, *PRMT7*, and *PRMT8*) were used in the subsequent analyses.

### Genes Associated with PRMT-RGs in DN

The enrichment scores of PRMT-RGs in the DN group were significantly higher than those in the control group. Therefore, the enrichment scores of PRMT-RGs were treated as traits for WGCNA (Figure 2C).

The data of all samples in the GSE30122 dataset were clustered, and no outliers were detected (Figure 2D). To ensure biologically meaningful scale-free topology, 17 was chosen as the minimal  $\beta$  value based on scale independence of  $> 0.8$ . (Figure 2E and F). In total, 15 co-expression modules were identified based on Pearson correlation analysis (Figure 2G). The pink module ( $R = -0.650$ ,  $p = 0.001$ ) comprising 643 genes exhibited the highest correlation with PRMT-RG scores (Figure 2H).

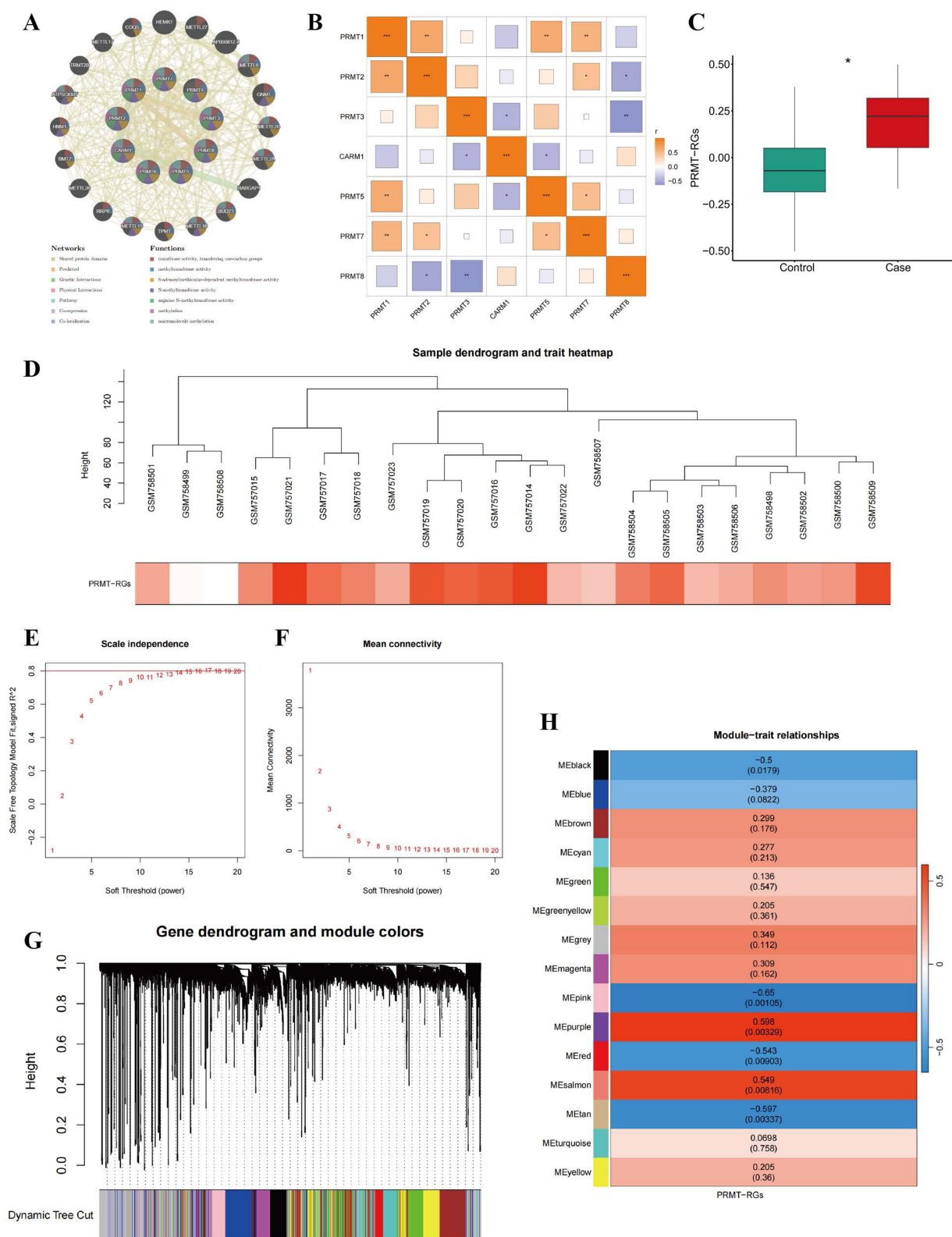
### DEGs Were Strongly Associated with Immune-Related Functions and Cell Adhesion

The data of the DN and control groups were subjected to differential gene analysis. In the training set, 716 DEGs (622 upregulated genes and 94 downregulated genes) were identified (Figure 3A and B).

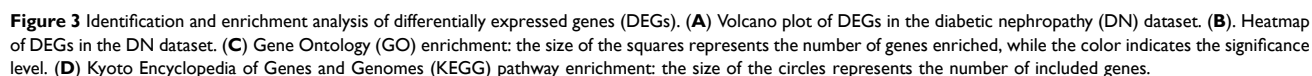
Biological functions and signaling pathways of the DEGs were analyzed using GO and KEGG enrichment analyses. GO analysis revealed that 716 DEGs were enriched in 970 GO items ( $\text{adj.}p < 0.05$ ) (Supplementary Table 1), which were related to positive regulation of cell-cell adhesion, positive regulation of cell adhesion, leukocyte-mediated immunity, positive regulation of T cell activation, and positive regulation of leukocyte cell-cell adhesion (Figure 3C). Meanwhile, the DEGs were enriched in 48 KEGG signaling pathways ( $\text{adj.}p < 0.05$ ) (Supplementary Table 2), including the phagosome, viral myocarditis, cell adhesion molecule, tuberculosis, and leishmaniasis pathways (Figure 3D).

### FAM13B, FAM98A, HPS5, PLBD1, and TGFBR3 Served as Candidate Genes

The number of intersection genes among 716 DEGs and 643 module genes was 15 (Figure 4A). Further, the 5 feature genes (*FAM13B*, *FAM98A*, *HPS5*, *PLBD1*, and *TGFBR3*) were selected using LASSO analysis after minimizing the model error ( $\lambda = 0.1$ ) (Figure 4B and C). Subsequently, the SVM-RFE model was constructed for the training dataset to screen the feature



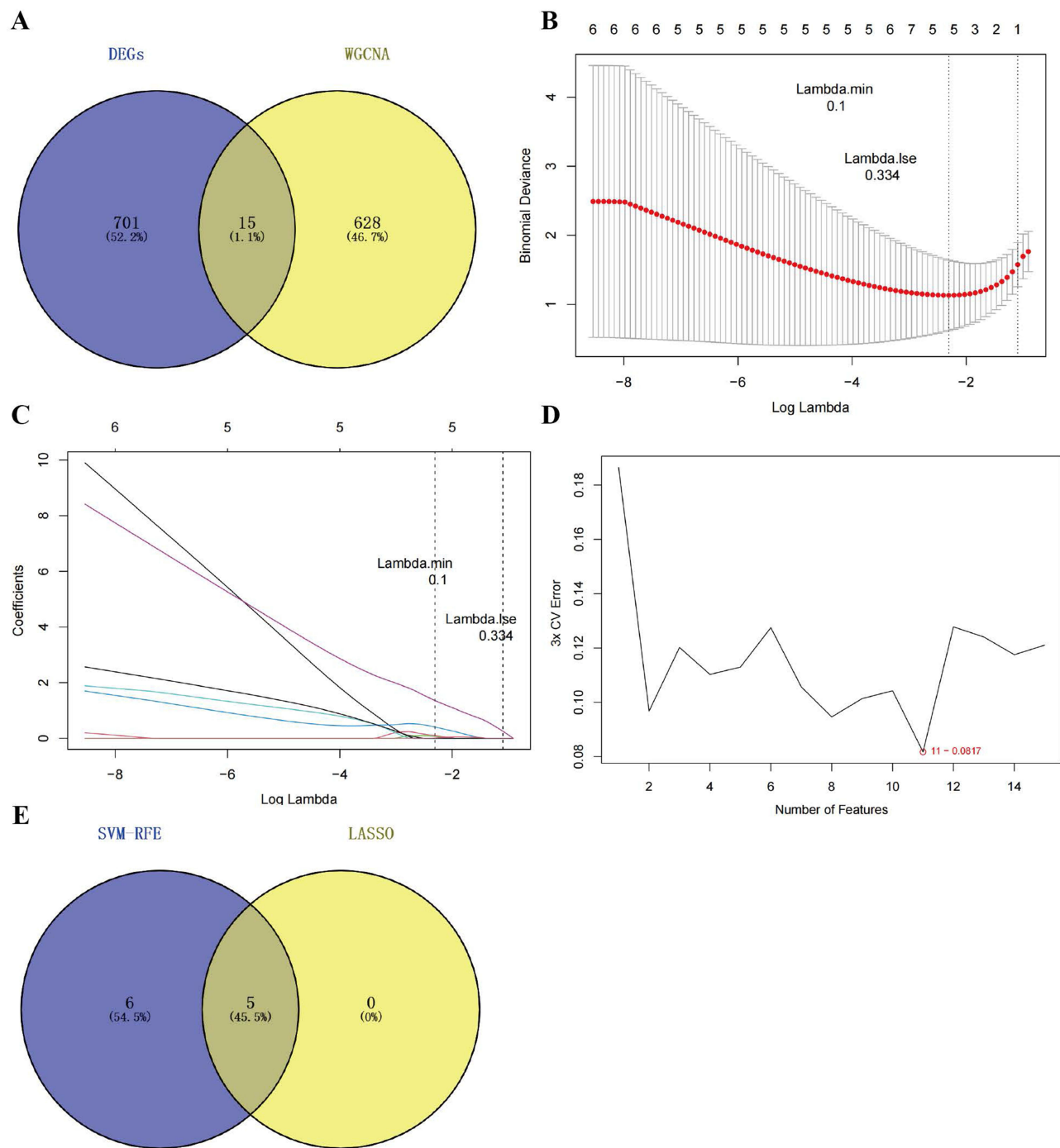
**Figure 2** Correlation analysis between protein arginine methylation-related genes (PRMT-RGs). **(A)** Gene-gene interaction (GGI) network of PRMT-RGs. **(B)** Spearman correlation coefficient analysis of PRMT-RGs. **(C)** The differential enrichment scores of PRMT-RGs between the diabetic nephropathy (DN) and control groups. **(D)** Sample dendrogram and trait heatmap (the branches represent samples, while the vertical axis represents the height of hierarchical clustering). **(E and F)** A scale-free topological model determines the optimal  $\beta$  value.  $\beta = 17$  was selected as the soft threshold based on average connectivity and scale independence. **(G)** Hierarchical clustering dendrogram of module identifiers. The color-coded rows below the dendrogram indicate module assignments determined by dynamic tree cuts. **(H)** A graph of the correlation between the module genes and PRMT-RGs. Rows and columns correspond to module eigengenes and traits, respectively.



## Enhanced Diagnostic Efficacy of FAM13B and FAM98A for DN

The diagnostic efficacy of the 5 candidate genes was evaluated using the ROC curve in the training and validation datasets. The area under the curve (AUC) values for *FAM13B*, *FAM98A*, *HPS5*, *PLBD1*, and *TGFBR3* in the training dataset were 0.958, 0.967, 0.933, 0.983, and 0.958, respectively, while those in the validation set were 0.921, 0.921, 0.675, 0.675, and 0.976, respectively. Therefore, *FAM13B*, *FAM98A*, and *TGFBR3* exhibited a good diagnostic efficacy for DN (AUC > 0.7) (Figure 5A and B). Differential expression analysis revealed that *FAM13B*, *FAM98A*, and *TGFBR3* were differentially expressed. In particular, *FAM13B* and *FAM98A* were significantly upregulated in the validation and training datasets. Consequently, *FAM13B* and *FAM98A* were identified as potential biomarkers for DN (Figure 5C and D). The nomogram demonstrated the contribution of *FAM13B* and *FAM98A* to the incidence of DN. DCA suggested the enhanced benefit and clinical utility of the nomogram (Figure 5E and F).



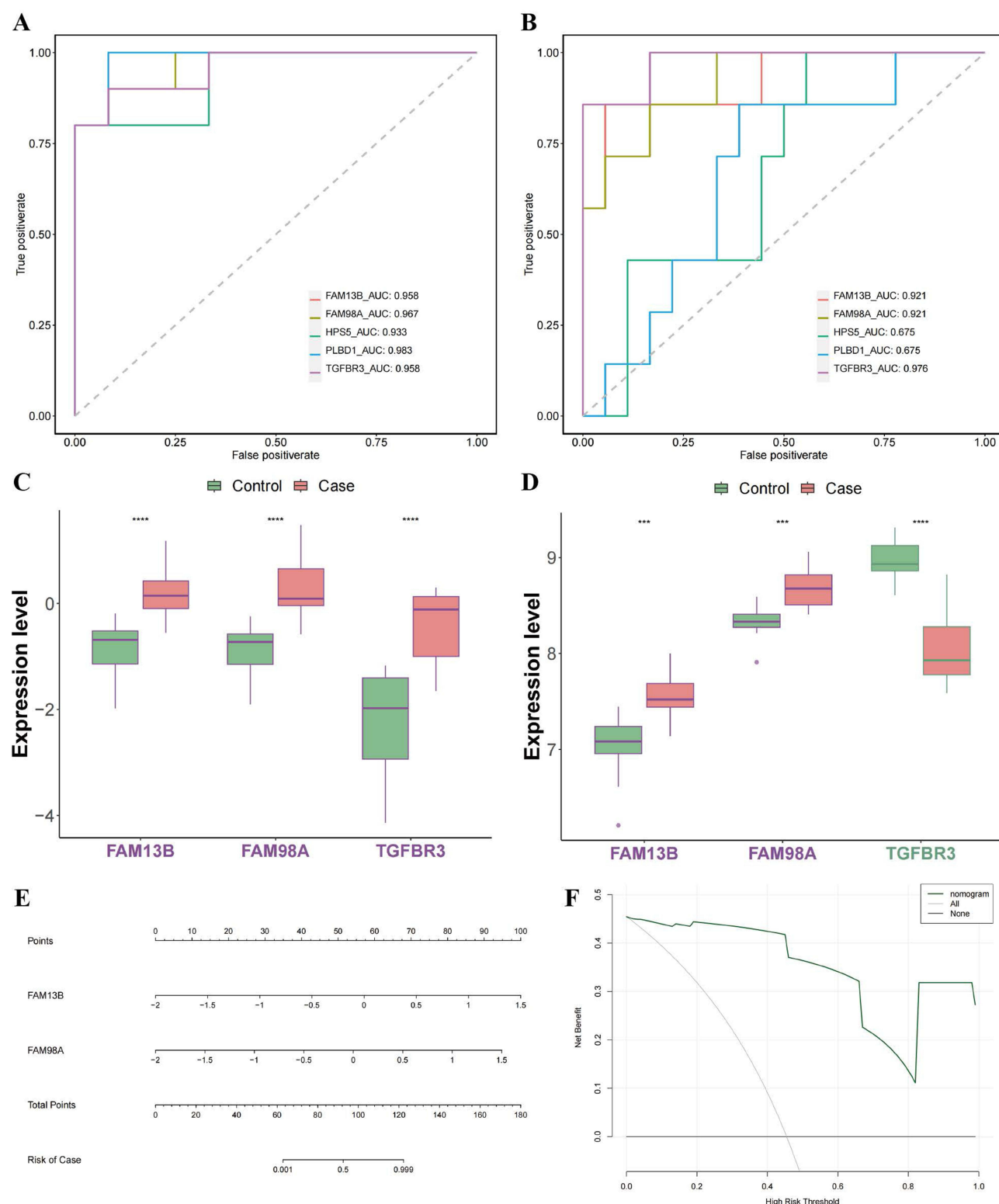


**Figure 4** Identification of candidate genes for diabetic nephropathy (DN) using least absolute shrinkage and selection operator (LASSO) and support vector machine-recursive feature elimination (SVM-RFE) machine learning methods. **(A)** Venn diagram: In total, 15 candidate genes were obtained after intersecting the 716 differentially expressed genes (DEGs) with the 643 module genes. **(B and C)** The LASSO logistic regression algorithm was used to identify five candidate genes of DN. **(D)** The SVM-RFE model was employed to identify 11 feature genes. **(E)** Venn diagram: Five target genes were obtained after intersecting genes obtained from SVM-RFE with those obtained from LASSO analyses.

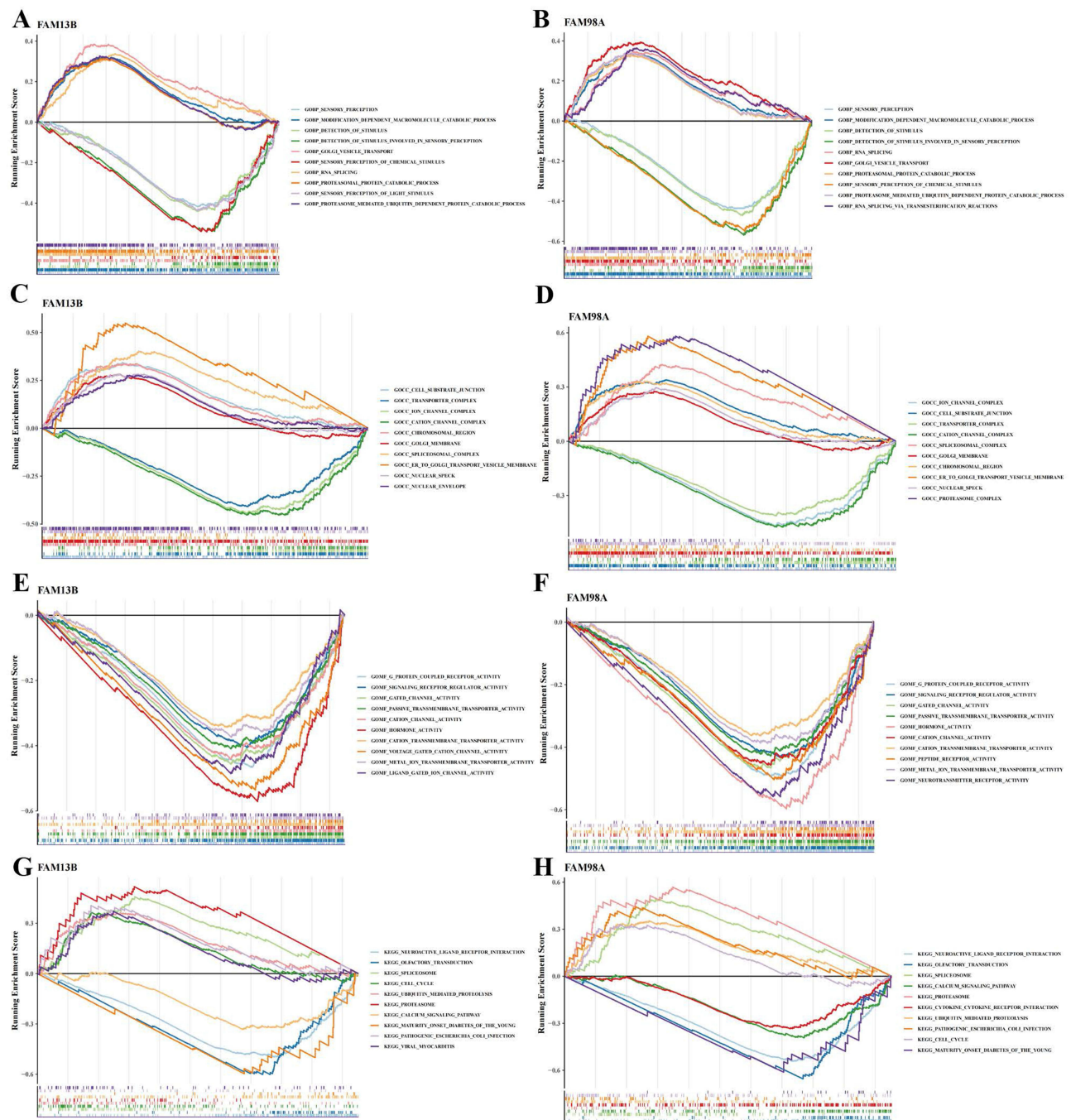
## FAM13B and FAM98A were Mainly Enriched in Sensory Perception and Modification-Dependent Macromolecule Catabolic Process

GSEA revealed that genes related to *FAM13B* and *FAM98A* were mainly enriched in sensory perception, modification-dependent macromolecule catabolic process, ion channel complex, G-protein-coupled receptor activity, neuroactive ligand-receptor interaction, and olfactory transduction signaling pathways (Figure 6A–H).





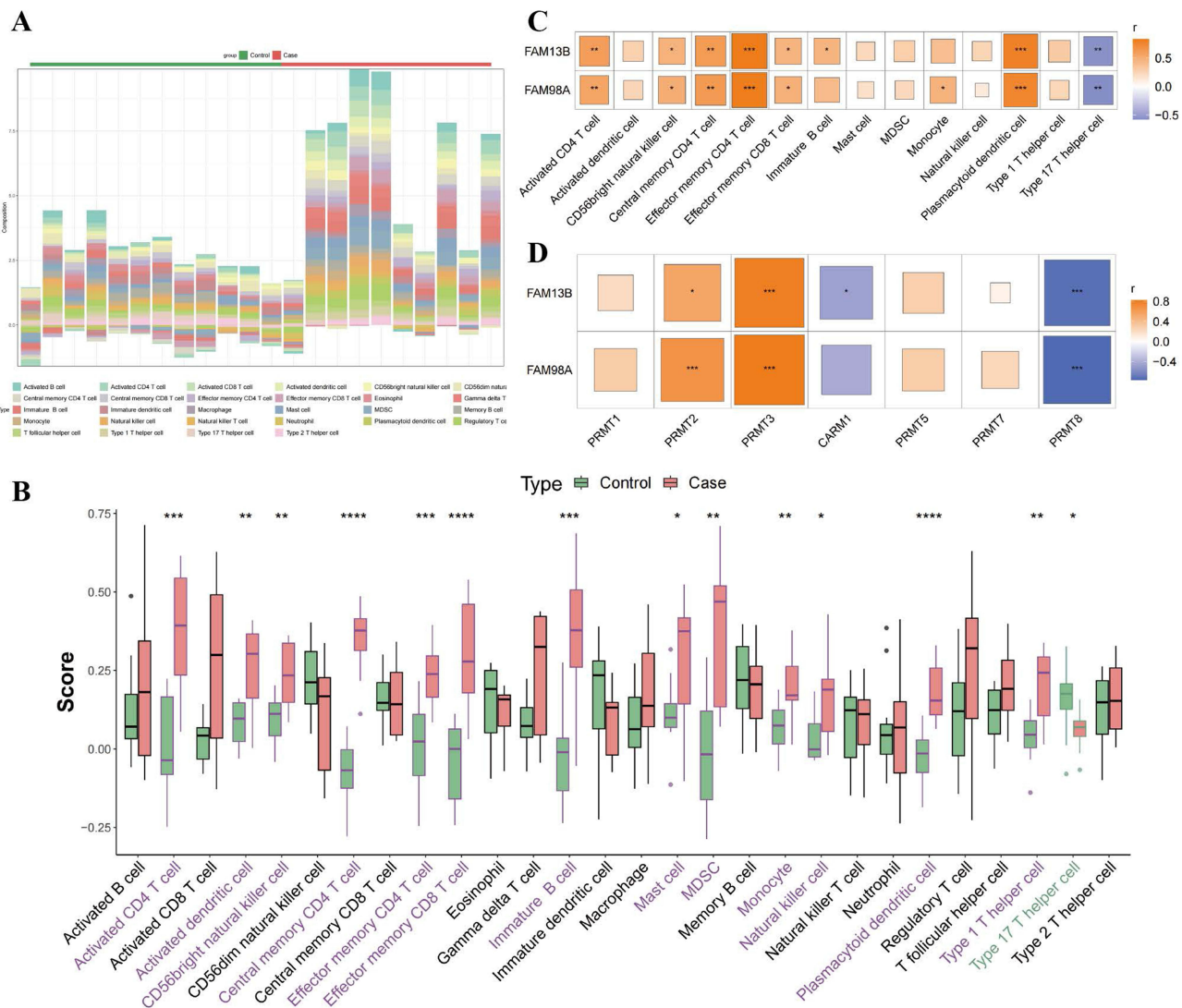
**Figure 5** Diagnostic nomogram model construction and efficacy assessment. **(A and B)** Receiver operating characteristic (ROC) profiles of five candidate genes between the control and diabetic nephropathy (DN) groups in the training set GSE30122 **(A)** and the validation set GSE104954 **(B)**. **(C and D)** The expression levels of three candidate biomarkers in the training set GSE30122 **(C)** and the validation set GSE104954 **(D)**. **(E)** The construction of the nomogram model based on FAM13B and FAM98A. \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . **(F)** Decision curve analysis (DCA) of the nomogram model.



**Figure 6** Gene set enrichment analysis (GSEA) of signaling pathways in which *FAM13B* and *FAM98A* are enriched. (A, C, E, and G). The enrichment analysis of *FAM13B*. (B, D, F, and H). The enrichment analysis of *FAM98A*.

## Th17 Cells are Involved in DN Pathogenesis

The infiltration status of 28 immune cells in the control and DN groups was illustrated using heatmaps (Figure 7A). Compared with that in the control group, the proportion of 14 immune cells was markedly altered in the DN group. Meanwhile, compared with those in the control group, the proportions of activated CD4 T cells, central memory CD4 T cells, effector memory CD4 T cells, effector memory CD8 T cells, immature B cells, and plasmacytoid were significantly upregulated and the proportions of type 17 T helper cell (Th17 cell) were downregulated in the DN group (Figure 7B). Analysis of the correlation between biomarkers and different immune cells revealed that *FAM13B* expression was positively correlated with effector memory CD4



**Figure 7** Immune infiltration analysis. **(A)** Heatmap depicting the infiltration pattern of immune cells. **(B)** The differential immune cell enrichment scores between the control and diabetic nephropathy (DN) groups were analyzed using the Wilcoxon rank-sum test. **(C)** Heatmap depicting the correlation between biomarkers and differentially enriched immune cells. **(D)** Heatmap illustrating the correlation between biomarkers and protein arginine methylation-related genes (PRMT-RGs). \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001; \*\*\*\*p < 0.0001.

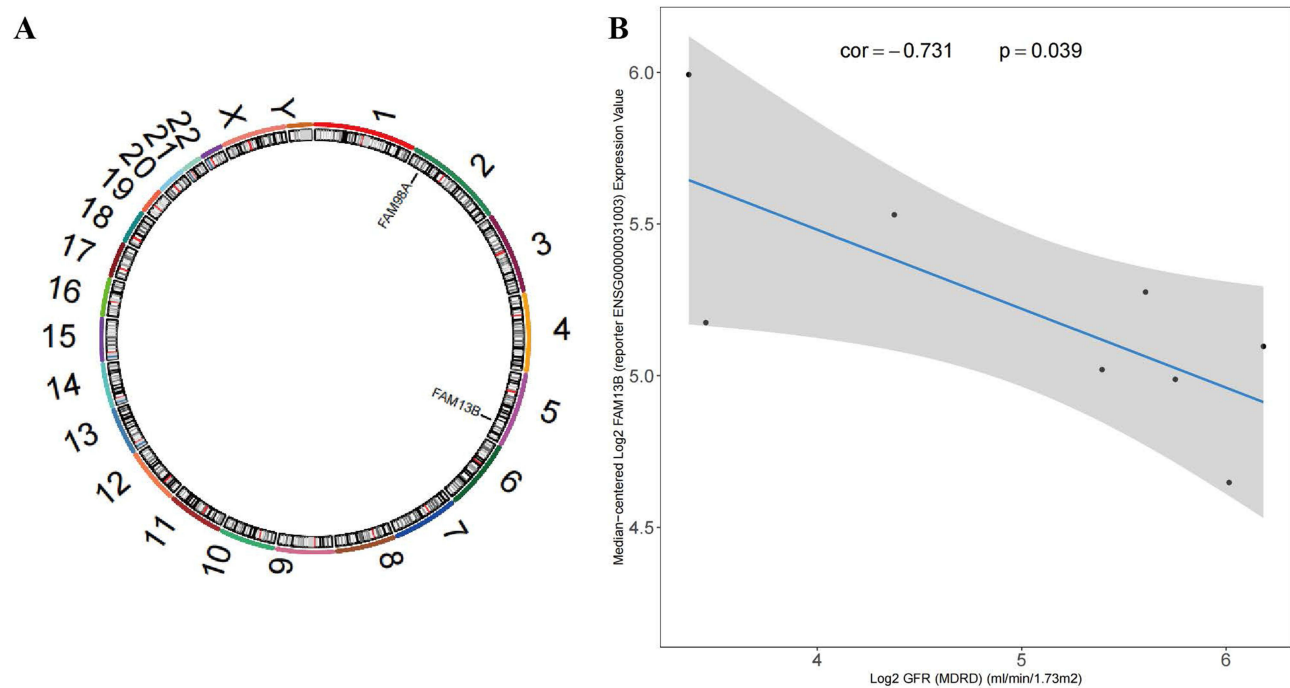
T cell proportion ( $R = 0.828$ ). In contrast, *FAM98A* expression was significantly and negatively correlated with Th17 cell proportion ( $R = -0.575$ ). (Figure 7C). Meanwhile, analysis of the correlation between biomarkers and PRMT-RGs revealed that *FAM98A* expression was significantly and positively correlated with *PRMT3* expression ( $R = 0.824$ ) but was significantly and negatively correlated with *PRMT8* expression ( $R = -0.764$ ) (Figure 7D).

## FAM13B Expression is Negatively Correlated with GFR in DN

*FAM13B* and *FAM98A* were located on chromosomes 5 and 2 of the human genome, respectively (Figure 8A). Additionally, *FAM13B* expression was negatively correlated with GFR in DN ( $\text{Cor} = -0.731$ ,  $p = 0.039$ ) (Figure 8B).

## FAM98A and FAM13B are Co-Regulated by Related Molecules

The biomarker-miRNA regulatory network predicted that 61 and 63 miRNAs target *FAM98A* and *FAM13B*, respectively. Of these, *FAM98A* and *FAM13B* were co-regulated by 6 miRNAs (hsa-mir-224-5p, hsa-mir-200b-3p, hsa-mir-124-3p, hsa-mir-1-3p, hsa-mir-218-5p, and hsa-mir-136-5p) (Figure 9A and B). The TF-biomarker regulatory network indicated that *FAM98A* and *FAM13B* were regulated by 17 and 14 TFs, respectively. TGIF2 regulated both *FAM98A* and *FAM13B* (Figure 9C).



**Figure 8** Chromosomal localization and clinical analysis of *FAM98A* and *FAM13B*. **(A)** Chromosomal localization: *FAM13B* is located on the fifth chromosome of the human genome, while *FAM98A* is located on the second chromosome. **(B)** Correlation between *FAM13B* expression and glomerular filtration rate (GFR).

## Estradiol and Rotenone are Potential Therapeutics for DN

The biomarker-drug network predicted that 40 and 35 drugs target *FAM98A* and *FAM13B*, respectively. Of these, seven drugs (ICG 001, estradiol, bisphenol A, rotenone, sodium arsenite, 1,2-dimethylhydrazine, and 7,8-dihydro-7,8-dihydroxybenzo(a)pyrene 9,10-oxide) targeted both *FAM98A* and *FAM13B* (Figure 9D). *FAM98A* exhibited favorable docking affinities with estradiol (binding energy =  $-5.24$  kcal/mol) and rotenone (binding energy =  $-5.36$  kcal/mol) (Table 2). Estradiol and amino acids (ARG-301, ARG-287, and SER-282) could bind to *FAM98A* through hydrogen bonds. In contrast, *FAM98A* did not interact with rotenone through hydrogen bonds but through hydrophobic interactions,  $\pi$ - $\pi$  stacking, cation- $\pi$ , or other non-hydrogen bonding interactions (Figure 9E and F).

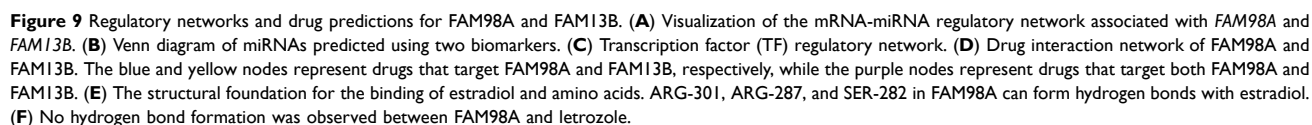
## Validation of Biomarker Expression in the DN Mouse Model

To validate the expression of biomarkers in biological samples, a DN mouse model was constructed. The successful establishment of the model was confirmed by monitoring blood glucose levels, body weight changes, and kidney tissue staining with MASSON and PAS methods (Supplementary Tables 3, 4 and Supplementary Figure 1). qRT-PCR analysis revealed that the expression levels of *Fam98a* and *Fam13b* were significantly upregulated in the experimental group (Figure 10A and B). These findings were consistent with those of GSE30122 and GSE104954 dataset analyses.

## FAM98A Expression Elevated in the Renal Tubular Epithelial Cells of DN Patients

To further validate the expression of the *FAM98A* protein in biological samples, we collected kidney paraffin-embedded tissue sections from patients undergoing renal biopsy at our center. Immunohistochemical staining revealed that *FAM98A* is primarily expressed in human renal tubular epithelial cells, and its expression is significantly increased in DN patients compared to patients with MCD (Figure 10C). These findings also were consistent with those of GSE30122 and GSE104954 dataset analyses.





Previous studies have reported that PRMTs are involved in the development and maintenance of DN. However, the regulatory mechanisms or the potential biomarkers of DN have not been elucidated. This study used a combination of differential expression analysis and diverse machine learning approaches to identify two biomarkers of DN (*FAM13B* and *FAM98A*). Comprehensive bioinformatics analysis and validation using mouse kidney samples revealed that *FAM13B* and *FAM98A* exhibited a robust diagnostic performance for DN. *FAM13B* and *FAM98A* were significantly correlated with *PRMT3*, *PRMT8*, effector memory CD4 T cells, and Th17 cells, contributing to the onset and progression of DN. Additionally, a regulatory network for the biomarkers was constructed, which enhanced our understanding of the molecular mechanisms underlying DN. Finally, drugs targeting the biomarkers were predicted, offering potential therapeutic options for the clinical management of DN.

Diabetes, Metabolic Syndrome and Obesity 2024:17

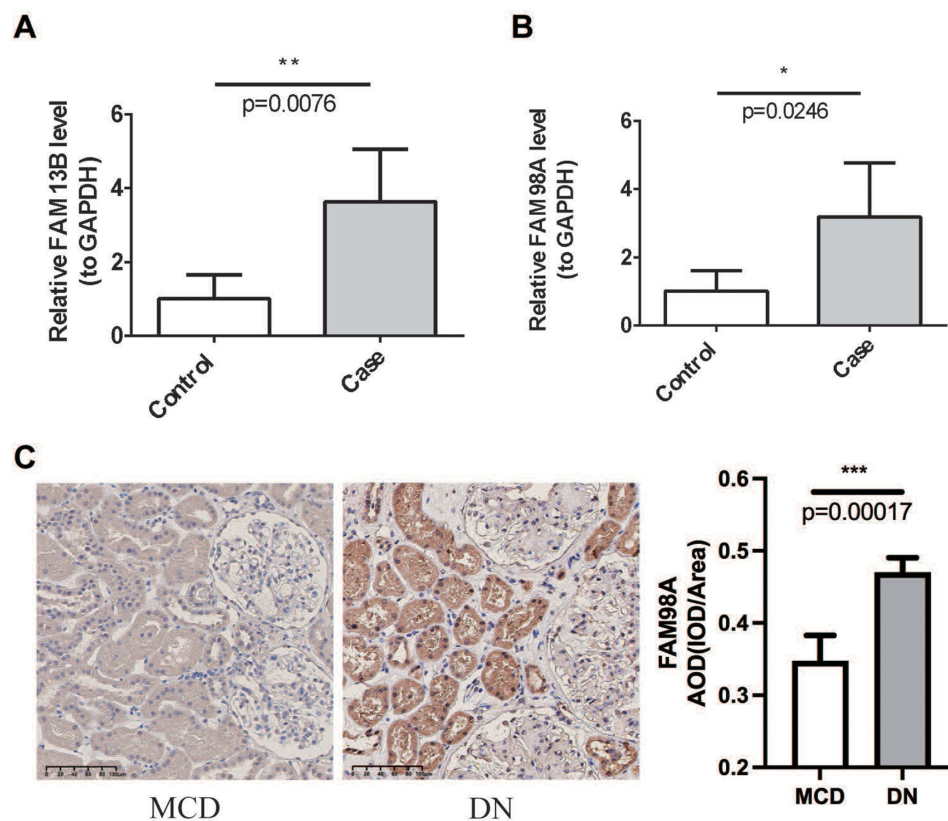


**Table 2** Binding Energy of Biomarker Binding to Targeted Drugs

Biomarkers	Targeted Drugs	Compound CID	Binding Energy (kcal/mol)
FAM98A	Bisphenol A	6623	-4.31
	7,8-Dihydro-7,8-dihydroxybenzo(a)pyrene 9.10-oxide	41233	-4.93
	Estradiol	5757	-5.24
	ICG 001	11,238,147	-3.93
	Rotenone	6758	-5.36
FAM13B	Bisphenol A	6623	-3.88
	7,8-Dihydro-7,8-dihydroxybenzo(a)pyrene 9.10-oxide	41233	-4.83
	Estradiol	5757	-4.93
	ICG 001	11,238,147	-3.3
	Rotenone	6758	-4.86

**Abbreviations:** DN, diabetic nephropathy; PRMT-RGs, protein arginine methylation-related genes; TFs, transcription factors; GFR, glomerular filtration rate; PRMT, protein arginine methylation; AKI, acute kidney injury; DEGs, differential expressed genes; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GGI, Gene-gene interaction network; WGCNA, weighted gene co-expression network analysis; ssGSEA, single sample Gene Set Enrichment Analysis; TOM, topological overlap matrix; LASSO, least absolute shrinkage and selection operator; SVM-RFE, support vector machine-recursive feature elimination; ROC, receiver operating characteristic; DCA, decision curve analysis; GSEA, gene set enrichment analysis; CTD, Comparative Toxicogenomics Database; STZ, streptozotocin; qRT-PCR, quantitative real-time polymerase chain reaction; cDNA, complementary DNA; AUC, area under curve; Th17 cell, type 17 T helper cell; TGF- $\beta$ , transforming growth factor- $\beta$ ; TNF- $\alpha$ , tumor necrosis factor- $\alpha$ ; IL-1 $\beta$ , interleukin-1 $\beta$ ; NF- $\kappa$ B, nuclear factor- $\kappa$ B; ERS, endoplasmic reticulum stress.

lipid metabolism in renal tubular epithelial cells and consequently alleviating DN has not been investigated. Among the five candidate genes, *TGFBR3* is critical for diagnosing glomerular lesions in DN.<sup>42</sup> Meanwhile, clinical evidence suggests that inflammation is a key etiological factor in DN.<sup>43–47</sup> *TGFBR3* specifically binds to transforming growth



**Figure 10** Validation of the expression patterns of *FAM13B* and *FAM98A*. (A) Quantitative real-time polymerase chain reaction (qRT-PCR) analysis revealed that the mRNA levels of *Fam13b* are upregulated in diabetic nephropathy (DN). (B) qRT-PCR analysis revealed that the mRNA levels of *Fam98a* are upregulated in DN. \* $p < 0.05$ ; \*\* $p < 0.01$ . (C) IHC illustrations and quantitative results of *FAM98A* in the kidneys of patients with MCD and patients with DN. \*\*\* $p < 0.001$ .

factor- $\beta$  (TGF- $\beta$ ), inducing the polarization of Th1 cells to M1 macrophages, which secrete tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), interleukin-1 $\beta$  (IL-1 $\beta$ ), and IL-12 to promote the inflammatory response of Th1 cells. In DN, the number of Th1 cells increases around the proximal tubular epithelial cells. Further, TNF- $\alpha$  induces hormone transcription and promotes inflammation and cell apoptosis through nuclear factor- $\kappa$ B (NF- $\kappa$ B) signal transduction.<sup>48,49</sup> *PLBD1*, a phospholipase mainly expressed in the bone marrow, is expressed in the kidneys.<sup>50</sup> Additionally, *PLBD1* functions as a lipid mediator and induces inflammation. The upregulation of *PLBD1* exacerbates heart failure after myocardial infarction, indicating that *PLBD1* expression may represent a general inflammatory signal.<sup>51</sup> However, the role of *PLBD1* in DN has not been reported. This study used SVM-RFE and LASSO analyses to demonstrate that *PLBD1* is a key gene involved in DN pathogenesis. The underlying molecular mechanism of *PLBD1* may involve the induction of inflammation, leading to the development of DN. These findings suggest that *HPS5*, *PLBD1*, and *TGFBR* are potential diagnostic genes for DN.

In this study, five genes identified from the training and validation sets were subjected to ROC curve analysis. *FAM13B*, *FAM98A*, and *TGFBR3* exhibited strong diagnostic ability (AUC values > 0.7). The expression levels of *FAM98A* and *FAM13B* were upregulated in the DN group. qRT-PCR analysis confirmed the upregulation of *Fam98a* and *Fam13b* in the kidneys of DN mice, which was consistent with the results of bioinformatics analysis. These findings suggest that *FAM98A* and *FAM13B* have diagnostic values in DN. Thus, *FAM98A* and *FAM13B* were considered biomarkers for DN.

The roles of *FAM98A* and *FAM13B* in the kidney and DN pathogenesis are unclear. Akter et al first identified *FAM98A* as a substrate for *PRMT1* using mass spectrometry. The C-terminus of proteins contains multiple RGG/RG motifs, which can be recognized by *PRMT1*, promoting direct arginine methylation.<sup>52</sup> Additionally, *Prmt1* is upregulated in the kidneys of DN mice,<sup>53</sup> exerting regulatory effects through the induction of oxidative stress, inflammation, and endoplasmic reticulum stress (ERS). In the STZ-induced DN rat model, *Prmt1* expression is upregulated. The suppression of *Prmt1* upregulation effectively improved the histopathological changes in the kidneys of diabetic rats.<sup>54</sup> This study performed Spearman correlation analysis to examine the correlation of *FAM13B* and *FAM98A* with PRMT-RGs and demonstrated that *FAM98A* and *FAM13B* were positively correlated with *PRMT3*. This suggests that both *FAM98A* and *FAM13B* are potential substrates for *PRMT3*, which is a novel discovery. In the invasive micropapillary carcinoma model, *PRMT3* regulates the ERS signaling pathway by promoting the arginine methylation of substrates.<sup>55</sup> Additionally, *PRMT3* is involved in activating ERS and mediating cisplatin-induced ototoxicity processes.<sup>56</sup> However, the post-translational modifications and folding of proteins typically occur in the endoplasmic reticulum and Golgi apparatus.<sup>57</sup> In this process, vesicular transport from the endoplasmic reticulum to the Golgi membrane plays a crucial role. GSEA revealed that *FAM98A* was positively associated with Golgi vesicle transport and proteasomal-mediated ubiquitin-dependent protein catabolic process. In the term cellular components, both *FAM98A* and *FAM13B* were positively correlated with Golgi membrane and endoplasmic reticulum to Golgi transport vesicle membrane. The results of this study and previous studies suggest that the overexpression of *FAM98A*, a substrate of *PRMT1* in DN renal tubular epithelial cells, may lead to accelerated cell death by inducing ERS through *PRMT1*-mediated arginine methylation, resulting in kidney damage. Further in vivo and in vitro experiments to elucidate the underlying mechanisms are currently in the planning stages.

Inflammation and immunity play crucial roles in the pathogenesis of DN. Preclinical studies have demonstrated the roles of inflammation and immunity in the pathogenesis of DN and the progression of renal damage.<sup>58,59</sup> Previous studies have identified the presence of effector memory CD4 T cells in the urine of patients with DN.<sup>60</sup> However, the infiltration status of effector memory CD4 T cells in the kidneys in DN has not been previously elucidated. This study identified two biomarkers. ssGSEA revealed the infiltration of effector memory CD4 T cells in the kidneys of patients with DN. *FAM13B* and *FAM98A* exhibited the strongest positive correlation with effector memory CD4 T cells. This indicates that the upregulation of *FAM98A* and *FAM13B* is critical for the inflammation and damage mediated by effector memory CD4 T cells in DN. Further experimental studies are needed to understand the specific contributions of CD4 T cells to the pathophysiology of DN.

This study demonstrates innovation in research methodology by successfully identifying biomarkers for DN through the use of public databases and bioinformatics techniques. Importantly, this study did not limit analysis to the theoretical level but also conducted experimental validation using animal and patient samples to confirm the existence and functionality of these biomarkers, thereby ensuring the reliability and practicality of the findings. Additionally, this study offers new insights into the

pathogenesis of DN by thoroughly examining the correlation between two key biomarkers—arginine methyltransferase and immune cells. Finally, the study predicted potential drugs targeting these biomarkers, which may provide new avenues for developing molecular-targeted therapy strategies for DN. In conclusion, this study not only offers novel perspectives on the diagnosis and clinical treatment of DN but also showcases unique innovation in both research methodology and the application of results. However, this study has not yet fully elucidated the molecular mechanisms underlying the biomarkers of diabetic nephropathy (DN), and it is crucial to validate these findings using a larger number of biological samples from DN patients. To address this, we plan to design and conduct more complex in vitro and in vivo validation experiments. In the future, we will thoroughly investigate the specificity, sensitivity, and clinical applicability of the newly identified biomarkers FAM98A and FAM13B, with the aim of advancing early diagnosis and intervention for DN. Additionally, we intend to first explore the regulatory effects of potential pharmacological compounds on FAM98A and FAM13B through in vitro studies, followed by assessing their impact on the progression of DN in animal models. Moreover, we will expand our research scope to explore the association of these biomarkers with other pathophysiological mechanisms of DN and actively search for new therapeutic targets.

## Conclusion

This study performed comprehensive bioinformatics analysis to identify 5 DEGs in DN (*FAM13B*, *FAM98A*, *HPS5*, *PLBD1*, and *TGFBR*). ROC curve analysis and expression validation in two databases revealed that *FAM98A* and *FAM13B* serve as biomarkers with an AUC value of  $> 0.9$  and exhibit consistent expression patterns. Further studies revealed that these biomarkers were correlated with PRMT3, PRMT8, effector memory CD4 T cell, and Th17 cell, contributing to DN onset and progression. qRT-PCR analysis revealed that the expression levels of *Fam98a* and *Fam13b* were significantly upregulated in the kidneys of DN mice. Additionally, the CTD database was used to predict potential drugs that target FAM98A and FAM13B, offering novel therapeutic modalities for DN.

## Data Sharing Statement

The datasets (GSE30122 and GSE104954) used and analyzed in the current study are available from GEO database (<https://www.ncbi.nlm.nih.gov/gds>).

## Ethics Approval and Consent to Participate

The animal protocol conformed to the guidelines outlined in “Laboratory Animal- Guidelines for Ethical Review of animal Welfare (GB/T 35892-2018)” to ensure laboratory animals’ ethical treatment and welfare. All procedures involving animals were approved by the ethical committee of Capital Medical University (Permit Number: AEEI-2022-284), Beijing, China. All procedures involving human participants and human data including publicly available data were approved by the ethical committee of Beijing Friendship Hospital, Capital Medical University (Permit Number: BFH20240528005/BFHHZS20240146), Beijing, China.

## Acknowledgments

Thanks all the contributors to the GEO database.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; 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 sponsored by Department of Nephrology at the Beijing Friendship Hospital, Capital Medical University (to X Wang), the National Natural Science Foundation of China (82003833, to L Wang), Beijing Friendship Hospital the Seed Program (YYZZ202114, to Y Guan).

## Disclosure

The authors declare that they have no competing interests.

## References

1. Yao L, Liang X, Qiao Y, Chen B, Wang P, Liu Z. Mitochondrial dysfunction in diabetic tubulopathy. *Metabolism*. 2022;131:155195. doi:10.1016/j.metabol.2022.155195
2. Vasanth Rao VR, Tan SH, Candasamy M, Bhattamisra SK. Diabetic nephropathy: an update on pathogenesis and drug development. *Diabetes Metab Syndr*. 2019;13(1):754–762. doi:10.1016/j.dsx.2018.11.054
3. de Boer IH, de Boer IH, Rue TC, et al. Temporal trends in the prevalence of diabetic kidney disease in the United States. *JAMA*. 2011;305(24):2532–2539. doi:10.1001/jama.2011.861
4. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. *N Engl J Med*. 2019;380(24):2295–2306. doi:10.1056/NEJMoa1811744
5. Heerspink H, Langkilde AM, Wheeler DC. Dapagliflozin in patients with chronic kidney disease. *Reply N Engl J Med*. 2021;384(4):389–390. doi:10.1056/NEJMc2032809
6. Pitt B, Filippatos G, Agarwal R, et al. Cardiovascular events with finerenone in kidney disease and type 2 diabetes. *N Engl J Med*. 2021;385(24):2252–2263. doi:10.1056/NEJMoa2110956
7. Bakris GL, Agarwal R, Anker SD, et al. Effect of finerenone on chronic kidney disease outcomes in type 2 diabetes. *N Engl J Med*. 2020;383(23):2219–2229. doi:10.1056/NEJMoa2025845
8. Herrington WG, Baigent C, Haynes R. Empagliflozin in patients with chronic kidney disease. *N Engl J Med*. 2023;388(24):2301–2302. doi:10.1056/NEJMc2301923
9. Wu Q, Schapira M, Arrowsmith CH, Barsyte-Lovejoy D. Protein arginine methylation: from enigmatic functions to therapeutic targeting. *Nat Rev Drug Discov*. 2021;20(7):509–530. doi:10.1038/s41573-021-00159-8
10. Zhu Y, Yu C, Zhuang S. Protein arginine methyltransferase 1 mediates renal fibroblast activation and fibrogenesis through activation of Smad3 signaling. *Am J Physiol Renal Physiol*. 2020;318(2):F375–F387. doi:10.1152/ajprenal.00487.2019
11. Nakayama Y, Ueda S, Yamagishi S, et al. Asymmetric dimethylarginine accumulates in the kidney during ischemia/reperfusion injury. *Kidney Int*. 2014;85(3):570–578. doi:10.1038/ki.2013.398
12. Mei X, Zeng J, Liu DF, et al. Abnormalities of the PRMT1-ADMA-DDAH1 metabolism axis and probucol treatment in diabetic patients and diabetic rats. *Ann Palliat Med*. 2021;10(3):3343–3353. doi:10.21037/apm-21-417
13. Park MJ, Han HJ, Kim DI. Lipotoxicity-induced PRMT1 exacerbates mesangial cell apoptosis via endoplasmic reticulum stress. *Int J Mol Sci*. 2017;18(7):1421. doi:10.3390/ijms18071421
14. Akalin PK. Introduction to bioinformatics. *Mol Nutr Food Res*. 2006;50(7):610–619. doi:10.1002/mnfr.200500273
15. Rauschert S, Raubenheimer K, Melton PE, Huang RC. Machine learning and clinical epigenetics: a review of challenges for diagnosis and classification. *Clin Clin Epigenet*. 2020;12(1):51. doi:10.1186/s13148-020-00842-4
16. Basu S, Faghmous JH, Doupe P. Machine learning methods for precision medicine research designed to reduce health disparities: a structured tutorial. *Ethn Dis*. 2020;30(Suppl 1):217–228. doi:10.18865/ed.30.S1.217
17. Terranova N, Venkatakrishnan K, Benincosa LJ. Application of machine learning in translational medicine: current status and future opportunities. *AAPS J*. 2021;23(4):74. doi:10.1208/s12248-021-00593-x
18. Pruneski JA, Williams RJ 3rd, Nwachukwu BU, et al. The development and deployment of machine learning models. *Knee Surg Sports Traumatol Arthrosc*. 2022;30(12):3917–3923. doi:10.1007/s00167-022-07155-4
19. Han H, Chen Y, Yang H, et al. Identification and verification of diagnostic biomarkers for glomerular injury in diabetic nephropathy based on machine learning algorithms. *Front Endocrinol*. 2022;13:876960. doi:10.3389/fendo.2022.876960
20. Na J, Sweetwyne MT, Park AS, Susztak K, Cagan RL. Diet-induced podocyte dysfunction in drosophila and mammals. *Cell Rep*. 2015;12(4):636–647. doi:10.1016/j.celrep.2015.06.056
21. Grayson PC, Eddy S, Taroni JN, et al. Metabolic pathways and immunometabolism in rare kidney diseases. *Ann Rheum Dis*. 2018;77(8):1226–1233. doi:10.1136/annrheumdis-2017-212935
22. Wei HH, Fan XJ, Hu Y, et al. A systematic survey of PRMT interactomes reveals the key roles of arginine methylation in the global control of RNA splicing and translation. *Sci Bull*. 2021;66(13):1342–1357. doi:10.1016/j.scib.2021.01.004
23. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;9:559. doi:10.1186/1471-2105-9-559
24. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
25. Cao T, Li Q, Huang Y, Li A. plotnineSeqSuite: a Python package for visualizing sequence data using ggplot2 style. *BMC Genomics*. 2023;24(1):585. doi:10.1186/s12864-023-09677-8
26. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849. doi:10.1093/bioinformatics/btw313
27. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2(3):100141. doi:10.1016/j.xinn.2021.100141
28. Walter W, Sánchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*. 2015;31(17):2912–2914. doi:10.1093/bioinformatics/btv300
29. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1–22. doi:10.18637/jss.v033.i01
30. Li Z, Qin Y, Liu X, et al. Identification of predictors for neurological outcome after cardiac arrest in peripheral blood mononuclear cells through integrated bioinformatics analysis and machine learning. *Funct Integr Genomics*. 2023;23(2):83. doi:10.1007/s10142-023-01016-0



31. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf.* 2011;12:77. doi:10.1186/1471-2105-12-77
32. Wang M, Tang H, Chen X, et al. Opportunistic muscle evaluation during chest CT is associated with vertebral compression fractures in old adults: a longitudinal study. *J Gerontol a Biol Sci Med Sci.* 2024;79(2):glad162. doi:10.1093/gerona/glad162
33. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. *BMC Bioinf.* 2013;14:244. doi:10.1186/1471-2105-14-244
34. Dapprih AL, Derks LM, Holtmann M, Lange WG, Legenbauer T, Becker ES. Hostile and threatening interpretation biases in adolescent inpatients are specific to callous-unemotional traits and social anxiety. *Eur Child Adolesc Psychiatry.* 2023;33:1143–1150. doi:10.1007/s00787-023-02227-3
35. Zhang C, Zheng Y, Li X, Hu X, Qi F, Luo J. Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. *Ann Transl Med.* 2019;7(18):427. doi:10.21037/atm.2019.08.113
36. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18(1):248–262. doi:10.1016/j.celrep.2016.12.019
37. Wang L, Wang D, Yang L, et al. Cuproptosis related genes associated with Jab1 shapes tumor microenvironment and pharmacological profile in nasopharyngeal carcinoma. *Front Immunol.* 2022;13:989286. doi:10.3389/fimmu.2022.989286
38. Wang X, Ning Y, Zhang P, Li C, Zhou R, Guo X. Hair multi-bioelement profile of Kashin-Beck disease in the endemic regions of China. *J Trace Elem Med Biol.* 2019;54:79–97. doi:10.1016/j.jtemb.2019.04.002
39. Zhang MM, Wang D, Lu F, et al. Identification of the active substances and mechanisms of ginger for the treatment of colon cancer based on network pharmacology and molecular docking. *BioData Min.* 2021;14(1):1. doi:10.1186/s13040-020-00232-9
40. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402–408. doi:10.1006/meth.2001.1262
41. Li J, Luo M, Hu M, et al. investigating the molecular mechanism of aqueous extract of cyclocarya paliurus on ameliorating diabetes by transcriptome profiling. *Front Pharmacol.* 2018;9:912. doi:10.3389/fphar.2018.00912
42. Li C, Su F, Zhang L, et al. Identifying potential diagnostic genes for diabetic nephropathy based on hypoxia and immune status. *J Inflamm Res.* 2021;14:6871–6891. doi:10.2147/JIR.S341032
43. Li S, Jia Y, Xue M, et al. Inhibiting Rab27a in renal tubular epithelial cells attenuates the inflammation of diabetic kidney disease through the miR-26a-5p/CHAC1/NF- $\kappa$ B pathway. *Life Sci.* 2020;261:118347. doi:10.1016/j.lfs.2020.118347
44. Salti T, Khazim K, Haddad R, Campisi-Pinto S, Bar-Sela G, Cohen I. Glucose induces IL-1 $\alpha$ -dependent inflammation and extracellular matrix proteins expression and deposition in renal tubular epithelial cells in diabetic kidney disease. *Front Immunol.* 2020;11:1270. doi:10.3389/fimmu.2020.01270
45. Zhou Y, Ma XY, Han JY, et al. Metformin regulates inflammation and fibrosis in diabetic kidney disease through TNC/TLR4/NF- $\kappa$ B/miR-155-5p inflammatory loop. *World J Diabetes.* 2021;12(1):19–46. doi:10.4239/wjd.v12.i1.19
46. Cui X, Li Y, Yuan S, et al. Alpha-kinase1 promotes tubular injury and interstitial inflammation in diabetic nephropathy by canonical pyroptosis pathway. *Biol Res.* 2023;56(1):5. doi:10.1186/s40659-023-00416-7
47. Yang WX, Liu Y, Zhang SM, et al. Epac activation ameliorates tubulointerstitial inflammation in diabetic nephropathy. *Acta Pharmacol Sin.* 2022;43(3):659–671. doi:10.1038/s41401-021-00689-2
48. Yang M, Zhang Y, Ren J. Acetylation in cardiovascular diseases: molecular mechanisms and clinical implications. *Biochim Biophys Acta Mol Basis Dis.* 2020;1866(10):165836. doi:10.1016/j.bbdis.2020.165836
49. Zouhal H, Zare-Kookandeh N, Haghighi MM, et al. Physical activity and adipokine levels in individuals with type 2 diabetes: a literature review and practical applications. *Rev Endocr Metab Disord.* 2021;22(4):987–1011. doi:10.1007/s11154-021-09657-x
50. Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics.* 2014;13(2):397–406. doi:10.1074/mcp.M113.035600
51. Vanhaverbeke M, Vausort M, Veltman D, et al. Peripheral Blood RNA Levels of QSOX1 and PLBD1 are new independent predictors of left ventricular dysfunction after acute myocardial infarction. *Circ Genom Precis Med.* 2019;12(12):e002656. doi:10.1161/CIRCGEN.119.002656
52. Akter KA, Mansour MA, Hyodo T, Ito S, Hamaguchi M, Senga T. FAM98A is a novel substrate of PRMT1 required for tumor cell migration, invasion, and colony formation. *Tumour Biol.* 2016;37(4):4531–4539. doi:10.1007/s13277-015-4310-5
53. Matsuguma K, Ueda S, Yamagishi S, et al. Molecular mechanism for elevation of asymmetric dimethylarginine and its role for hypertension in chronic kidney disease. *J Am Soc Nephrol.* 2006;17(8):2176–2183. doi:10.1681/ASN.2005121379
54. Ojima A, Ishibashi Y, Matsui T, et al. Glucagon-like peptide-1 receptor agonist inhibits asymmetric dimethylarginine generation in the kidney of streptozotocin-induced diabetic rats by blocking advanced glycation end product-induced protein arginine methyltransferase-1 expression. *Am J Pathol.* 2013;182(1):132–141. doi:10.1016/j.ajpath.2012.09.016
55. Zhi R, Wu K, Zhang J, et al. PRMT3 regulates the progression of invasive micropapillary carcinoma of the breast. *Cancer Sci.* 2023;114(5):1912–1928. doi:10.1111/cas.15724
56. Deith M, Brodie JF. Predicting defaunation: accurately mapping bushmeat hunting pressure over large areas. *Proc Biol Sci.* 2020;287(1922):20192677. doi:10.1098/rspb.2019.2677
57. Wang Y, Jin M, Cheng CK, Li Q. Tubular injury in diabetic kidney disease: molecular mechanisms and potential therapeutic perspectives. *Front Endocrinol.* 2023;14:1238927. doi:10.3389/fendo.2023.1238927
58. Rayego-Mateos S, Morgado-Pascual JL, Opazo-Ríos L, et al. Pathogenic pathways and therapeutic approaches targeting inflammation in diabetic nephropathy. *Int J Mol Sci.* 2020;21(11):3798. doi:10.3390/ijms21113798
59. Rayego-Mateos S, Rodríguez-Diez RR, Fernández-Fernández B, et al. Targeting inflammation to treat diabetic kidney disease: the road to 2030. *Kidney Int.* 2023;103(2):282–296. doi:10.1016/j.kint.2022.10.030
60. Kopetschke K, Klocke J, Griebbach AS, et al. The cellular signature of urinary immune cells in Lupus nephritis: new insights into potential biomarkers. *Arthritis Res Ther.* 2015;17(1):94. doi:10.1186/s13075-015-0600-y



**Diabetes, Metabolic Syndrome and Obesity****Dovepress****Publish your work in this journal**

Diabetes, Metabolic Syndrome and Obesity is an international, peer-reviewed open-access journal committed to the rapid publication of the latest laboratory and clinical findings in the fields of diabetes, metabolic syndrome and obesity research. Original research, review, case reports, hypothesis formation, expert opinion and commentaries are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/diabetes-metabolic-syndrome-and-obesity-journal>