# REG1A and RUNX3 Are Potential Biomarkers for Predicting the Risk of Diabetic Kidney Disease

Xinyu Wang[1†], Han Wu[2†], Guangyan Yang[1†], Jiaqing Xiang[1], Lijiao Xiong[1], Li Zhao[3], Tingfeng Liao[1], Xinyue Zhao[4], Lin Kang[1,5*], Shu Yang[1,6*] and Zhen Liang[1,6*]

[1] Department of Geriatrics, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, China, [2] Department of Endocrinology, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, China, [3] Department of Health Management, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, China, [4] Department of Nephrology, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, China, [5] The Biobank of National Innovation Center for Advanced Medical Devices, Shenzhen People's Hospital, Shenzhen, China, [6] Shenzhen Clinical Research Center for Aging, Shenzhen, China

Diabetic kidney disease (DKD) is the leading cause of end-stage renal disease. Clinical features are traditionally used to predict DKD, yet with low diagnostic efficacy. Most of the recent biomarkers used to predict DKD are based on transcriptomics and metabolomics; however, they also should be used in combination with many other predictive indicators. The purpose of this study was thus to identify a simplified class of blood biomarkers capable of predicting the risk of developing DKD. The Gene Expression Omnibus database was screened for DKD biomarkers, and differentially expressed genes (DEGs) in human blood and kidney were identified *via* gene expression analysis and the Least Absolute Shrinkage and Selection Operator regression. A comparison of the area under the curve (AUC) profiles on multiple receiver operating characteristic curves of the DEGs in DKD and other renal diseases revealed that *REG1A* and *RUNX3* had the highest specificity for DKD diagnosis. The AUCs of the combined expression of *REG1A* and *RUNX3* in kidney (AUC = 0.929) and blood samples (AUC = 0.917) of DKD patients were similar to each other. The AUC of blood samples from DKD patients and healthy individuals obtained for external validation further demonstrated that *REG1A* combined with *RUNX3* had significant diagnostic efficacy (AUC=0.948). *REG1A* and *RUNX3* expression levels were found to be positively and negatively correlated with urinary albumin creatinine ratio and estimated glomerular filtration rate, respectively. Kaplan-Meier curves also revealed the potential of *REG1A* and *RUNX3* for predicting the risk of DKD. In conclusion, *REG1A* and *RUNX3* may serve as biomarkers for predicting the risk of developing DKD.

Keywords: diabetic kidney disease, biomarkers, diagnosis, gene expression omnibus, disease risk prediction

# INTRODUCTION

With the increasing incidence of diabetes mellitus (DM) worldwide, diabetic kidney disease (DKD) has become the leading cause of end-stage renal disease with a high mortality rate (1). Approximately one-quarter of people with DM end up developing DKD (2). DKD diagnosis is primarily based on renal pathology and/or clinical manifestations and involves the determination of glomerular filtration rate (eGFR) and urinary protein levels in addition to clinical features such as the duration of DM and the presence of diabetic retinopathy (3, 4). However, eGFR and albuminuria are deficient in terms of sensitivity and specificity, while renal puncture is not widely available due to its invasiveness. Therefore, markers associated with DKD risk need to be identified for successful diagnosis and prevention. A meta-analysis showed that urinary transferrin excretion rate was a good predictor of DKD occurrence (5). Kidney injury molecule 1 is a membrane protein expressed in the apical membrane of proximal tubule cells and is of great value in assessing the progression of DKD (6). After 12 years of follow-up in a cohort study of 628 patients with DM, plasma tumor necrosis factor levels were found to be associated with early decline in eGFR was strongly correlated (7). Older age, male sex, prolonged diabetes duration, hypertension, glycated hemoglobin levels (HbA$_{1c}$), and plasma triglyceride levels were identified as risk factors for proteinuria in patients with type 2 DM in a previous meta-analysis of 13 studies (8). Several biomarkers to assess DKD risk have been identified as a result of recent advancements in analytical techniques as well. Mayer et al. also identified nine serum markers associated with decreased GFR in Type 2 DM, including chitinase 3-like protein 1, growth hormone 1, hepatocyte growth factor, matrix metalloproteinase 2 (MMP2), MMP7, MMP8, MMP13, tyrosine kinase, and tumor necrosis factor receptor 1 (9). The surrogate markers for micro- and macro-vascular hard endpoints for innovative diabetes tools study further identified five serum markers (fibroblast growth factor 21, symmetric-to-asymmetric dimethylarginine ratio, β2-microglobulin, C16-acylcarnitine, and kidney injury molecule 1) with significant prediction potential for GFR decline in patients with type 2 DM (10). Although traditional clinical characteristics are easily determined, their prognostic value is limited. In general, serum biomarkers have a higher predictive efficacy than clinical characteristics; however, multiple biomarkers should be used in combination to achieve such efficacy. Simple biomarker combinations often lack specificity. Thus, widespread application of serum biomarkers is hindered.

Genome-wide transcriptome analysis has been widely used in the field of DKD to help gain insight into disease pathogenesis, molecular classification and identification of biomarkers (11). These gene expression matrices are included in the Gene Expression Omnibus (GEO) database to facilitate integration and reanalysis by researchers. The goal of this study was to establish a simplified class of blood biomarkers for predicting the risk of DKD. The study flow is illustrated in **Figure 1**.



**FIGURE 1** | Flowchart. DKD, diabetic kidney disease; CKD, Chronic kidney disease.
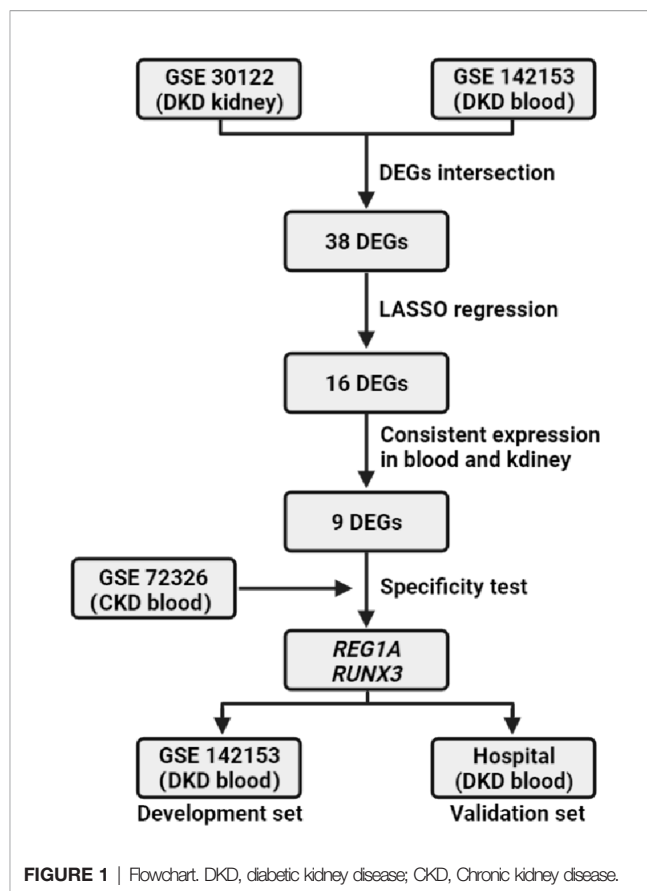
# METHOD

## Data Acquisition

We acquired gene expression profile datasets GSE30122 (11), GSE142153 (12), and GSE72326 (13) from the GEO database (14) (https://www.ncbi.nlm.nih.gov/geo/), and performed gene ID and symbol conversions *via* Perl scripts (15). GSE30122 included kidney tissue from patients with pathologically confirmed DKD stage 4-5, with healthy kidney tissue as a control. GSE142153 incorporates peripheral blood from patients with a clinical diagnosis of DKD stages 3-5, with healthy human blood as a control. GSE142153 included peripheral blood from patients with confirmed CKD, with blood from people without kidney disease as a control.

## Identification of Differentially Expressed Genes

R software (version 4.1.0, http://r-project.org/) was used for data analysis and plotting. The "limma" R package (16) was used to screen for DEGs, and heatmap and volcano plots of DEGs were constructed by the "ggplot2" package (17) to visualize the expression levels of DEGs. p < 0.05 was considered statistically significant.

## Screening and Validation of Diagnostic Markers

The Least Absolute Shrinkage and Selection Operator (LASSO) method, as implemented in the "glmnet" software package, was used for reduction of data dimensionality and selection of predictive features for DKD patients (18, 19). The receiver operating characteristic (ROC) curve method is used for evaluation of diagnostic performance (20). The area under the curve (AUC) of ROCs of single or multiple factors was calculated using the " pROC" software package (21). Calibration curves were plotted using the "rms" software package to assess whether the predicted probability of the model approximated the true probability (22, 23).

## Correlation of Diagnostic Markers With Clinical Characteristics

The "ggstatsplot" package was used to perform Spearman correlation analysis of diagnostic markers and clinical features (24), and the results were subsequently visualized using the "ggplot2" package.

## Analysis of the Prognostic Potential of Identified Biomarkers

Kaplan-Meier (KM) curves were constructed using "survival" and "survminer" software packages to assess the probability of DKD occurring at specific time periods, and log-rank tests were used to determine differences between groups (25). The prognostic value of the diagnostic markers was assessed using univariate or multivariate Cox proportional hazards models.

## Clinical Statistics

A total of 141 human blood samples from patients with DKD, DM (without DKD), and healthy individuals, were collected from the biospecimen bank of Shenzhen People's Hospital. The study was approved by the Ethics Committee of Shenzhen People's Hospital. Data are expressed as mean plus/minus standard deviation (SD) or median and interquartile range for continuous variables and as percentages for categorical variables. The Mann-Whitney U-test or t-test was performed to compare the differences between the two groups depending on whether the data conformed to a normal distribution. Chi-square test was applied to compare frequencies. This work was approved by the Institutional Review Board and the Ethics Committee of the Shenzhen People's Hospital (No. LL-KT-2018338) and informed written consent was obtained from all participants.

## Quantitative Real-Time PCR Analysis

Trizol (Invitrogen) was used to extract total RNA from peripheral blood mononuclear cells (PBMC) according to manufacturer's instructions. Reverse transcription of RNA was performed using the RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific). Quantitative PCR was performed using PowerUp SYBR Green Master Mix (Thermo Scientific). The results were standardized using GAPDH. qPCR was conducted using an ABI 7500 real-time PCR system (Applied Biosystems, Foster City, CA, USA). Fold-change was determined as $2^{-\triangle\triangle Ct}$ for gene expression. Gene-specific PCR primers are listed in **Table 1**.

# RESULTS

## Screening for DEGs Shared Across Blood and Kidney Samples

**Figure 2** shows the screening process for DEGs shared across blood and kidney biopsy samples from patients with DKD. Using the healthy control samples (HC) as basis, 679 (GSE142153: HC, n=10, DKD, n=23) and 499 (GSE30122: HC, n=50, DKD, n=19) DEGs were identified in blood and kidney samples from patients with DKD, respectively. A total of 38 DEGs were found to be shared across blood and kidney biopsy samples from patients with DKD.

## Screening for DEGs With Consistent Expression in DKD Blood and Kidney Samples

The LASSO logistic regression algorithm was used to further screen DEGs shared across blood samples from patients with DKD (GSE142153). A total of 16 DEGs were identified as diagnostic DKD markers (**Figures 3A, B**), including 5 and 11 down- and up-regulated genes, respectively (**Figure 3C**). Two of these DEGs were down-regulated whereas 14 were upregulated in kidney biopsy samples from patients with DKD (GSE30122) (**Figure 3D**). Ultimately, nine DEGs with consistent expression profiles in blood and kidney biopsy samples from patients with DKD were identified: which included Alpha-2A adrenergic receptor (*ADRA2A*), C-C motif chemokine 5 (*CCL5*), cholesterol 25-hydroxylase (*CH25H*), C-X-C chemokine receptor type 4 (*CXCR4*), hemoglobin subunit delta (*HBD*), hydroxycarboxylic acid receptor 3 (*HCAR3*), lysophosphatidylcholine acyltransferase 1 (*LPCAT1*), Lithostathine-1-alpha (*REG1A*), and Runt-related transcription factor 3 (*RUNX3*). The distribution of the expression profiles of these DEGs in the blood and kidney of DKD patients (**Figures 3E, F**) revealed that expressions of all of them were upregulated.

## Screening for Diagnostic DEGs in DKD

Two blood transcriptome datasets were analyzed to determine DKD-specific DEGs for diagnosis. In the GSE142153 dataset, several DEGs (*ADRA2A*, *CCL5*, *CH25H*, *CXCR4*, *HBD*, *HCAR3*, *LPCAT1*, *REG1A*, and *RUNX3*) were found to have high diagnostic efficiency in blood samples (AUC > 0.70), except for *CH25H* (AUC =0.67) (**Figure 4A**). The GSE72326 dataset [lupus nephritis (LN), n=48; ANCA-associated nephritis, n=10; focal segmental sclerosis (FSGS), n = 3; IgA nephropathy (IgAN), n = 5; microscopic disease (MCD), n =3] was used to validate the specificity of these DEGs for DKD diagnosis, except for *HCAR3*, which was excluded from the analysis as it was not found in the dataset. The diagnostic efficacies of eight DEGs in patients with LN were found to be relatively low (**Figure 4B**). *CXCR4*, *LPCAT1*, and *HBD* showed high diagnostic efficacy for ANCA-associated nephritis (**Figure 4C**). Significant diagnostic performance of *CXCR4*, *CCL5*, and *HBD* was observed for

**TABLE 1 |** The sequences of primers for qRT-PCR analysis.

| Gene | Forward | Backward |
| --- | --- | --- |
| *Homo RUNX3* (Gene ID: 864) | AGGCAATGACGAGAACTACTCC | CGAAGGTCGTTGAACCTGG |
| *Homo REG1A* (Gene ID: 5967) | ACCAGCTCATACTTCATGCTGA | CCAGGTCTCACGGTCTTCAT |
| *Homo GAPDH* (Gene ID: 2597) | GGAGCGAGATCCCTCCAAAAT | GGCTGTTGTCATACTTCTCATGG |

patients with FSGS (**Figure 4D**). *ADRA2A, LPCAT1,* and *HBD* showed high diagnostic efficacy for patients with IgAN (**Figure 4E**). *CH25H* and *HBD* showed good diagnostic efficacy in patients with MCD (**Figure 4F**). In summary, the transcription level of *REG1A* and *RUNX3* in blood samples may serve as DKD-specific predictors of disease development.

## DKD Diagnosis Efficacy of *RUNX3* and *REG1A* in Blood and Kidney Samples

Expression levels of *REG1A* and *RUNX3* were found to be significantly increased in blood samples of patients with DKD (**Figures 5A, B**; GSE142153). The combined diagnostic efficacy of these genes was also high (AUC=0.917, 95% CI: 0.818-1) (**Figure 5C**). Similar results have been obtained from kidney samples from DKD patients as well (**Figures 5D, E**; GSE30122, AUC=0.929, 95% CI: 0.846-1, **Figure 5F**). These findings indicate that *REG1A* and *RUNX3* have diagnostic potential for DKD in blood as well as kidney.

## Diagnostic Performance of *REG1A* and *RUNX3* in the Validation Set

A total of 141 blood samples from the human biospecimen bank of Shenzhen People's Hospital were used for qPCR analysis of *REG1A* and *RUNX3*. DKD (n = 50) and HC (n = 41) groups were included in the validation set (**Table 2**). The expression of *REG1A* and *RUNX3* were found to be significantly upregulated in the DKD group compared to those in the HC group (**Figures 6A, B**). The AUC for *REG1A* and *RUNX3* were found to be 0.912 and 0.859, respectively (**Figure 6C**). When *REG1A* and *RUNX3* were fitted as a single variable, the diagnostic efficiency was found to be 0.917 (**Figure 6C**) for the development set, and even higher for the validation set (AUC=0.948, 95% CI: 0.898-0.998) (**Figure 6D**), indicating that *REG1A* and *RUNX3* have high diagnostic value. A high degree of agreement was also found between the predicted and true values of the calibration curves in both development and validation sets (**Figures 6E, F**), indicating significant efficacy of *REG1A* and *RUNX3* for predicting DKD development.

## Analysis of Correlation Between Identified DEGs and Clinical Characteristics

The DKD (n=50) and DM (without DKD) groups (n=50) in the validation set were used to analyze correlation between expression levels of diagnostic DEGs and clinical characteristics (**Table 3**). Expression levels of *REG1A* and *RUNX3* were found to be significantly higher in the DKD group than those in the DM group (**Figures 7A, B**). *REG1A* was found to be positively correlated with serum creatinine (SCr), C-peptide (C-P), HbA$_{1C}$, fasting blood glucose (FBG), and urinary albumin creatinine ratio (UACR) and negatively correlated with eGFR level (**Figure 7C, Supplementary Figure 1**). *RUNX3* was positively correlated with UACR and SCr and negatively correlated with eGFR level (**Figure 7D, Supplementary Figure 2**). *REG1A* and *RUNX3* expression levels were found to be positively correlated as well (r=0.3) (**Supplementary Figure 1G**).

## KM Analysis of Diagnostic Markers and Clinical Characteristics

The KM method was used to analyze the probability of DKD occurrence in the corresponding time periods, and the variables included diagnostic DEGs and clinical characteristics. Using
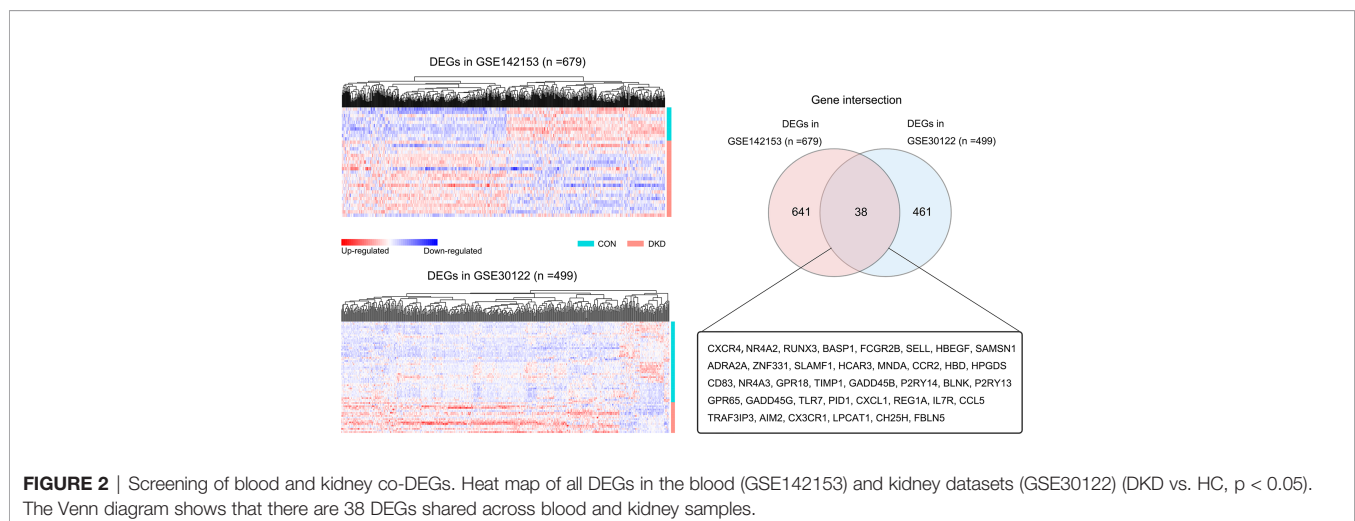


**FIGURE 2 |** Screening of blood and kidney co-DEGs. Heat map of all DEGs in the blood (GSE142153) and kidney datasets (GSE30122) (DKD vs. HC, p < 0.05). The Venn diagram shows that there are 38 DEGs shared across blood and kidney samples.
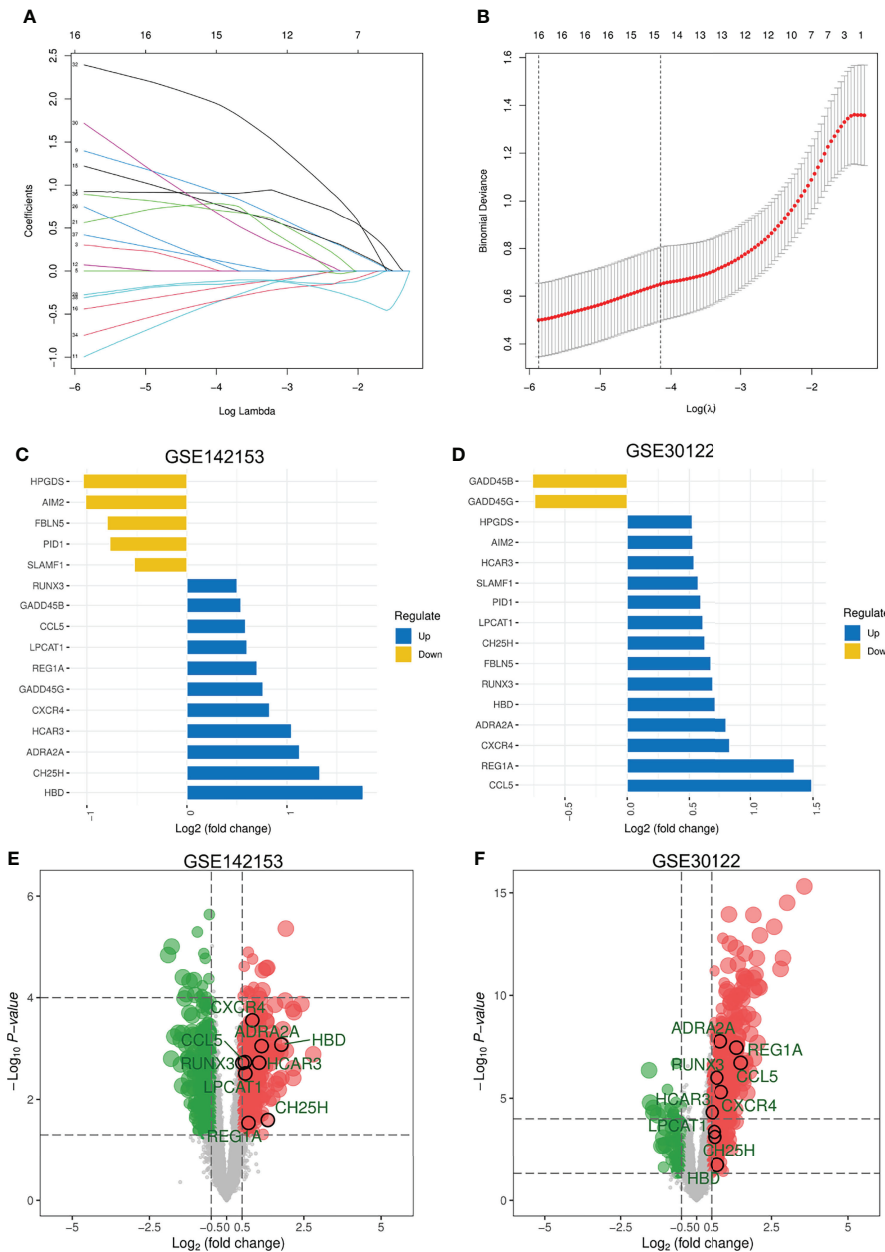
**FIGURE 3** | Screening for DEGs with Consistent Expression in DKD Blood and Kidney Samples. **(A)** LASSO coefficient profiles of the 38 features. A coefficient profile plot was produced against the log(λ) sequence. A vertical line was drawn at the value selected using five-fold cross-validation, where the optimal lambda resulted in 16 features with nonzero coefficients. **(B)** Optimal parameter (λ) selection in the LASSO model used five-fold cross-validation *via* minimum criteria. The partial likelihood deviance (binomial deviance) curve was plotted versus log(λ). Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 SE of the minimum criteria (the 1-SE criteria). **(C, D)** Expression of 16 DEGs in blood and kidney. Five DEGs that were consistently expressed in blood and kidney. **(E)** Distribution of these 5 DEGs in DKD blood samples. **(F)** Distribution of the 5 DEGs in the kidney of DKD. LASSO, least absolute shrinkage and selection operator; SE, standard error.

DKD as the endpoint event, variables associated with negative prognosis included *REG1A*, *RUNX3*, total cholesterol (TC), FBG, SCr, body mass index (BMI), and UACR, whereas elevated levels of high-density lipoprotein cholesterol (HDL-C), age, and eGFR indicated a positive prognosis (**Table 4** and **Supplementary Figure 3**). Combined analysis showed that patients with high expression of *REG1A* and *RUNX3* had the worst prognosis in all four groups (HR = 6.459) (**Figure 8**).
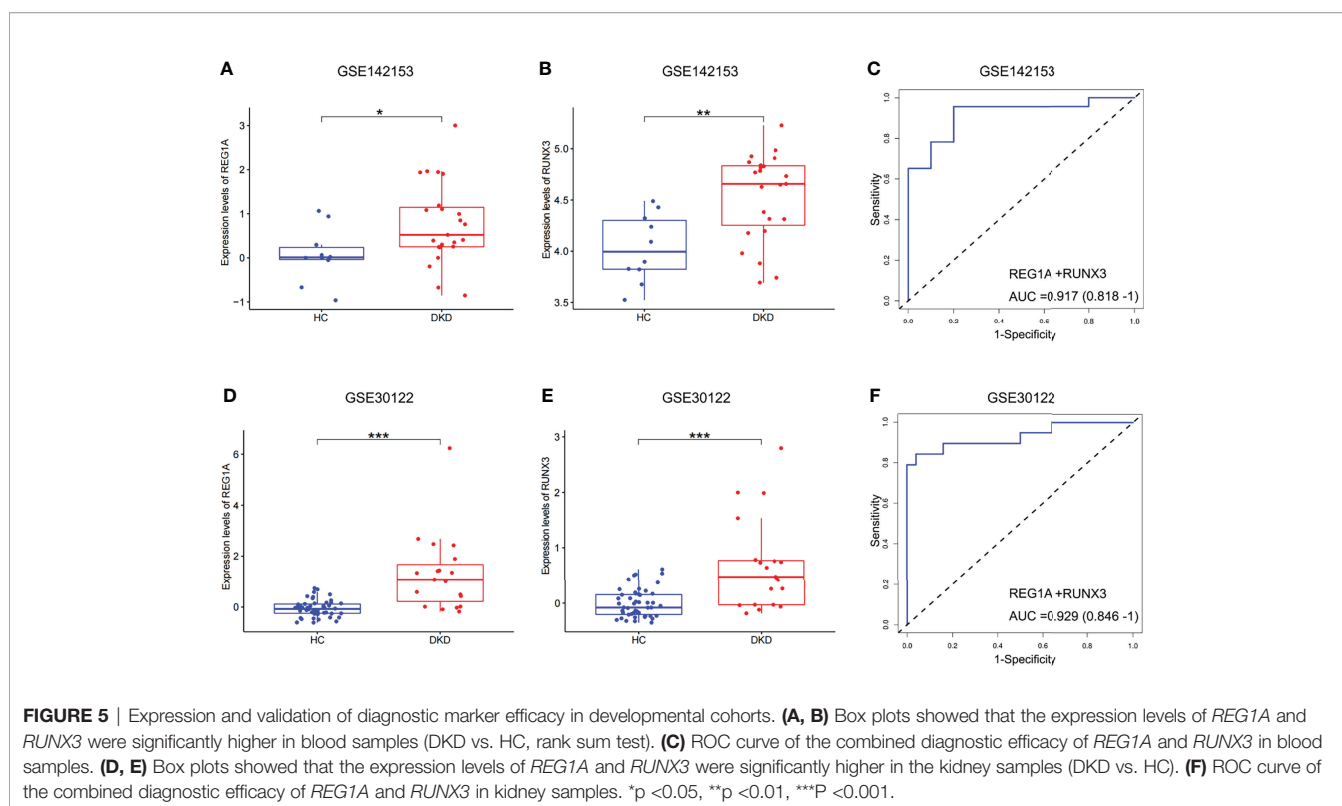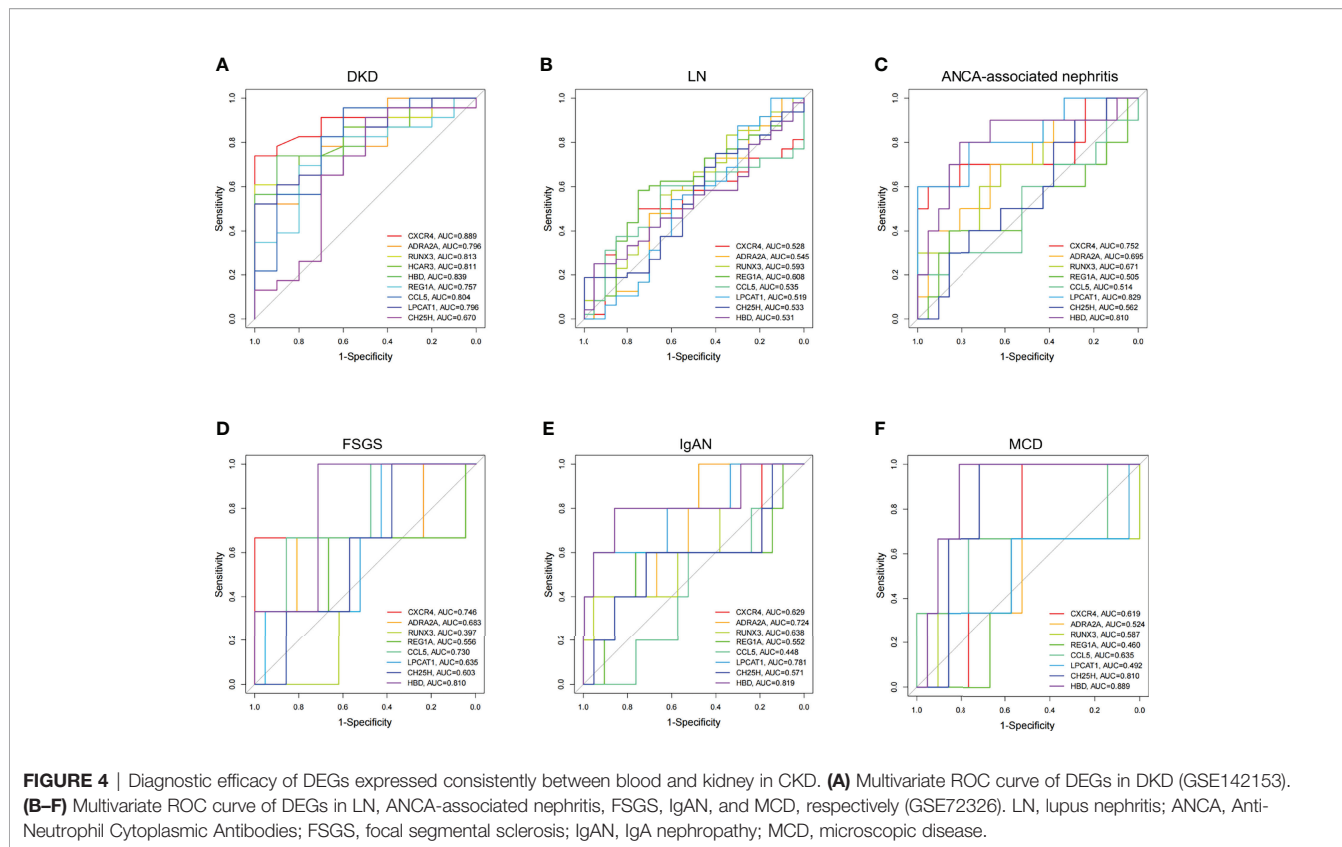
**FIGURE 4** | Diagnostic efficacy of DEGs expressed consistently between blood and kidney in CKD. **(A)** Multivariate ROC curve of DEGs in DKD (GSE142153). **(B–F)** Multivariate ROC curve of DEGs in LN, ANCA-associated nephritis, FSGS, IgAN, and MCD, respectively (GSE72326). LN, lupus nephritis; ANCA, Anti-Neutrophil Cytoplasmic Antibodies; FSGS, focal segmental sclerosis; IgAN, IgA nephropathy; MCD, microscopic disease.



**FIGURE 5** | Expression and validation of diagnostic marker efficacy in developmental cohorts. **(A, B)** Box plots showed that the expression levels of *REG1A* and *RUNX3* were significantly higher in blood samples (DKD vs. HC, rank sum test). **(C)** ROC curve of the combined diagnostic efficacy of *REG1A* and *RUNX3* in blood samples. **(D, E)** Box plots showed that the expression levels of *REG1A* and *RUNX3* were significantly higher in the kidney samples (DKD vs. HC). **(F)** ROC curve of the combined diagnostic efficacy of *REG1A* and *RUNX3* in kidney samples. *p <0.05, **p <0.01, ***P <0.001.

**TABLE 2 |** Baseline information on diagnostic markers and clinical characteristics in the validation cohort.

|  | Overall | HC | DKD | p |
|---|---|---|---|---|
| **n** | 91 | 41 | 50 | |
| **Female (%)** | 32 (35.2) | 17 (41.5) | 15 (30.0) | 0.358 |
| **Age (years)** | 56.00 [52.00, 66.00] | 54.00 [52.00, 61.00] | 57.00 [52.00, 67.00] | 0.211 |
| *RUNX3* | 1.49 [1.00, 2.82] | 1.07 [0.74, 1.38] | 2.63 [1.58, 3.66] | <0.001 |
| *REG1A* | 1.68 [0.94, 2.98] | 0.96 [0.84, 1.09] | 2.84 [2.02, 3.82] | <0.001 |

*Data are shown as mean (SD), median [25% quartile, 75% quartile] or numbers (percentages).*

## DISCUSSION

In the present study, we found that serum RUNX3 and REG1A have potential as diagnostic markers for DKD. Patients with DM who have elevated expression of both RUNX3 and REG1A will have a much higher risk of developing DKD at about 10 years into the disease. Compared to existing studies of diagnostic markers for DKD (26, 27), we have achieved high diagnostic efficacy with the combination of only two markers. Importantly, the diagnostic efficacy of RUNX3 and REG1A in blood is not inferior to that of renal tissue, which facilitates their widespread use. RUNX3 and REG1A are more specific in the diagnosis of
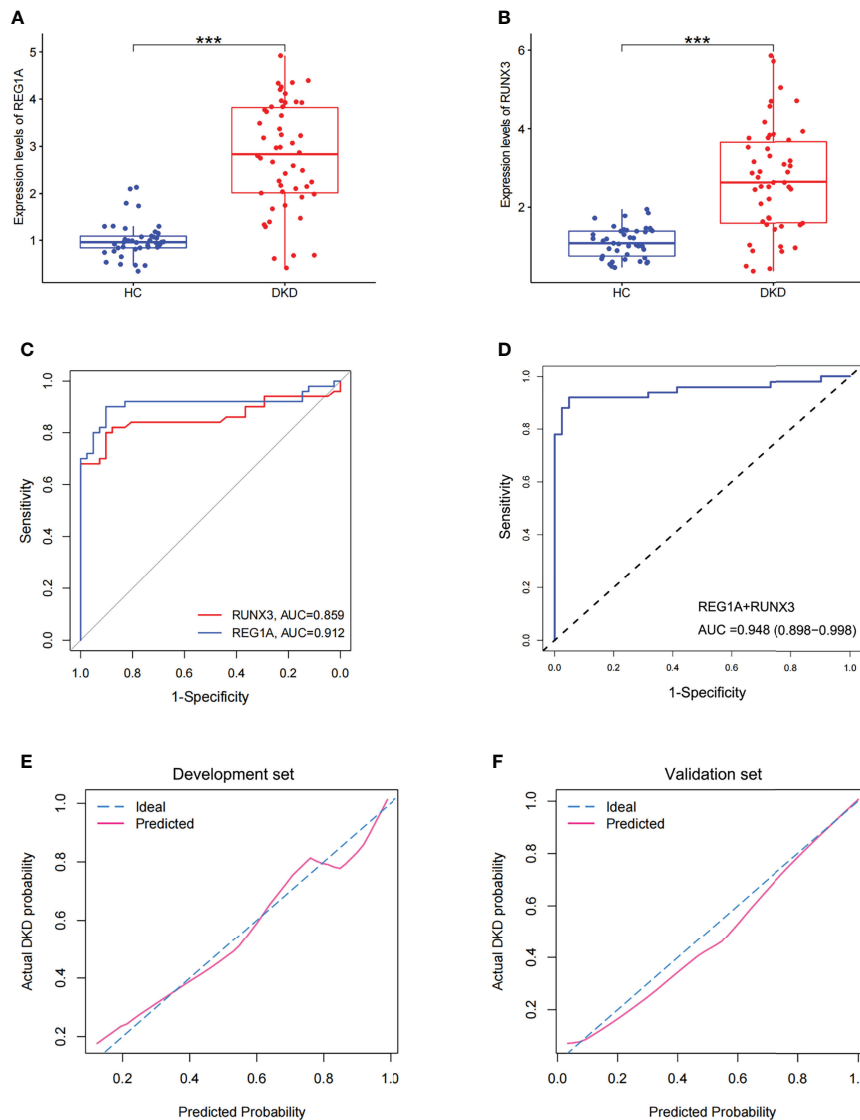


**FIGURE 6 |** Expression and diagnostic efficacy of diagnostic markers in the validation cohort. **(A, B)** Box plots showed significantly higher expression levels of *REG1A* and *RUNX3* (DKD vs. HC). **(C)** Separate ROC curves of *REG1A* and *RUNX3*. **(D)** The ROC curve of the diagnostic efficacy verification after fitting two diagnostic markers to one variable. **(E, F)** Calibration curves of diagnostic markers. The diagonal dotted line represents a perfect prediction by an ideal model. The solid line represents the performance of the nomogram, of which a closer fit to the diagonal dotted line represents a better prediction. The fit of both the dashed and solid lines for the development set GSE142153 **(E)** and the validation set **(F)** was excellent. ***P <0.001.

TABLE 3 | Baseline information on diagnostic markers and clinical characteristics in the DM and DKD groups.

| | Overall | DM | DKD | p |
|---|---|---|---|---|
| **n** | 100 | 50 | 50 | |
| **Female (%)** | 30 (30.0) | 15 (30.0) | 15 (30.0) | 1 |
| **Age (years)** | 57.65 (10.49) | 56.40 (9.63) | 58.90 (11.24) | 0.235 |
| **Duration of DM (years)** | 12.26 (6.59) | 10.59 (5.53) | 13.56 (7.10) | 0.016 |
| **BMI (kg/cm$^2$)** | 24.10 (3.15) | 24.13 (3.15) | 24.06 (3.18) | 0.916 |
| **FBG (mmol/L)** | 6.84 [5.54, 8.51] | 6.68 [5.61, 8.44] | 7.36 [5.46, 8.48] | 0.743 |
| **C-P (ng/ml)** | 1.65 [0.89, 2.51] | 1.65 [1.24, 2.30] | 1.58 [0.67, 3.07] | 0.986 |
| **HbA1c (%)** | 8.89 (2.09) | 8.54 (1.96) | 9.25 (2.17) | 0.092 |
| **SCr (μmol/L)** | 80.50 [65.00, 131.25] | 69.50 [60.25, 79.50] | 131.50 [86.75, 181.00] | <0.001 |
| **eGFR (ml/min/1.73 m$^2$)** | 84.21 [47.02, 101.26] | 98.13 [88.74, 104.92] | 46.30 [32.20, 80.08] | <0.001 |
| **UA (μmol/L)** | 375.22 (110.66) | 343.68 (97.29) | 406.76 (115.08) | 0.004 |
| **UACR (mg/g)** | 120.12 [7.42, 1932.66] | 7.40 [4.27, 16.40] | 1951.62 [1025.49, 2908.83] | <0.001 |
| **TG (mmol/L)** | 1.40 [0.97, 2.16] | 1.23 [0.92, 2.13] | 1.45 [1.10, 2.17] | 0.201 |
| **TC (mmol/L)** | 4.53 [3.62, 5.72] | 4.28 [3.63, 5.47] | 4.85 [3.68, 6.11] | 0.222 |
| **HDL-C (mmol/L)** | 1.07 [0.89, 1.21] | 1.08 [0.93, 1.22] | 1.05 [0.82, 1.21] | 0.218 |
| **LDL-C (mmol/L)** | 2.56 [1.85, 3.73] | 2.54 [1.88, 3.45] | 2.63 [1.85, 3.77] | 0.563 |
| *RUNX3* | 1.58 (1.07) | 1.00 (0.62) | 2.18 (1.10) | <0.001 |
| *REG1A* | 1.11 (0.47) | 1.00 (0.41) | 1.20 (0.49) | 0.02 |

*Data are shown as mean (SD), median [25% quartile, 75% quartile], or numbers (percentages). BMI, body mass index; FBG, fasting blood glucose; C-P, C-peptide; HbA1c, glycated hemoglobin A1c; SCr, serum creatinine; eGFR, estimation of glomerular filtration rate; UA, uric acid; UACR, urine albumin creatinine ratio; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoproteins cholesterol; LDL-C, low-density lipoproteins cholesterol.*

DKD compared to markers such as urinary transferrin, urinary IgG and urinary type IV collagen (28).

RUNX transcription factors regulate various biological processes, including embryonic development, cell proliferation, differentiation, cell lineage determination, and apoptosis (29).

*RUNX3* plays a downstream role in the TGF-β signalling pathway (30). Smyth et al. reported that *RUNX3* methylation is significantly increased in the blood of patients with DKD (31). In high glucose-treated renal tubular epithelial cells, RUNX3 was found to regulate the TGF-β1/Smad signalling pathway (32).



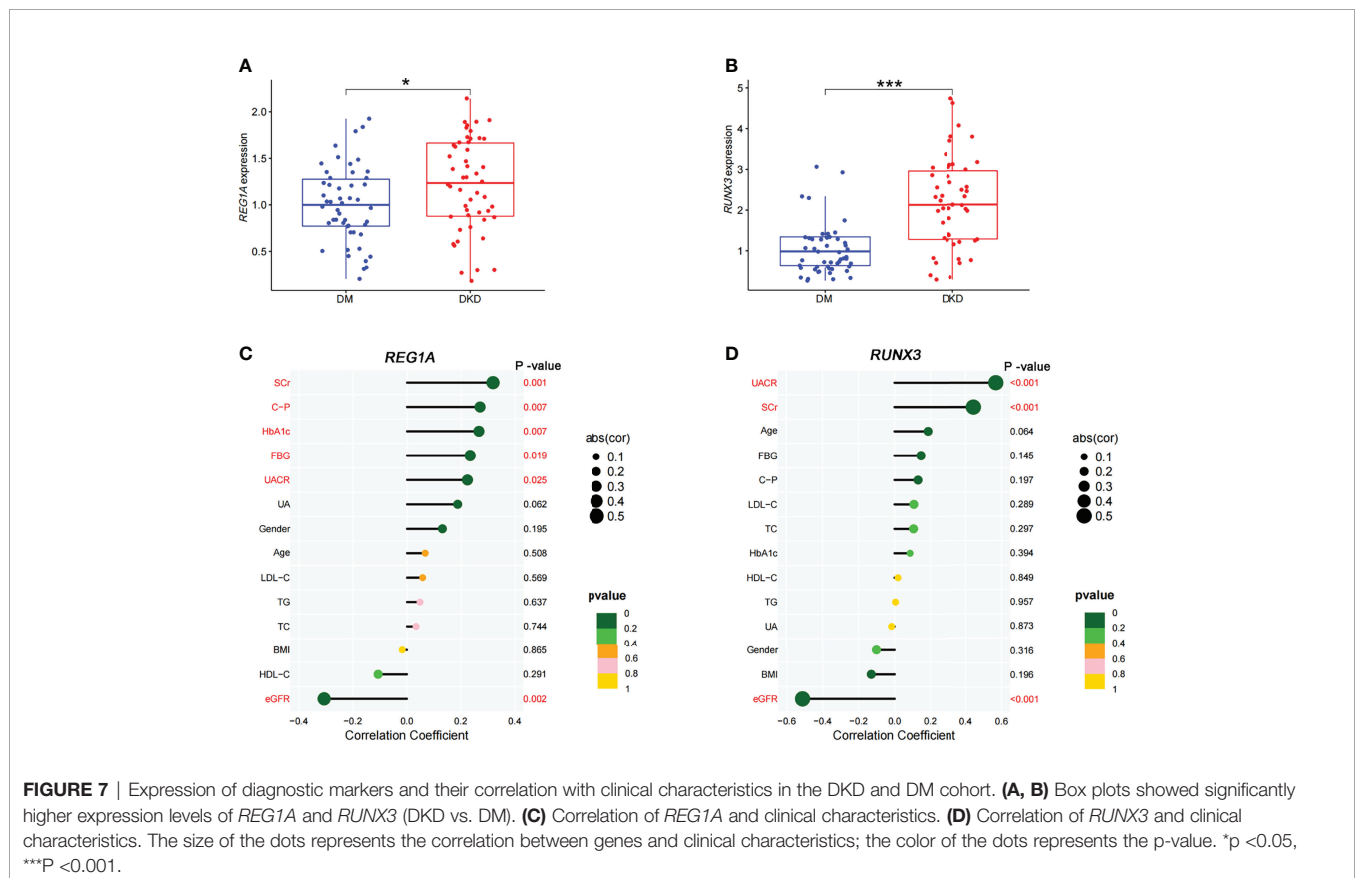FIGURE 7 | Expression of diagnostic markers and their correlation with clinical characteristics in the DKD and DM cohort. (A, B) Box plots showed significantly higher expression levels of *REG1A* and *RUNX3* (DKD vs. DM). (C) Correlation of *REG1A* and clinical characteristics. (D) Correlation of *RUNX3* and clinical characteristics. The size of the dots represents the correlation between genes and clinical characteristics; the color of the dots represents the p-value. *p <0.05, ***P <0.001.

**TABLE 4 |** Univariate COX regression of diagnostic markers and clinical characteristics.

| Variable | HR | lower 95%CI | upper 95%CI | p-value |
|---|---|---|---|---|
| *RUNX3* | 2.50 | 1.43 | 4.36 | <0.001 |
| *REG1A* | 2.97 | 1.35 | 4.17 | 0.003 |
| TC | 1.93 | 1.11 | 3.39 | 0.022 |
| FBG | 1.78 | 1.01 | 3.12 | 0.029 |
| SCr | 2.44 | 1.40 | 4.25 | 0.001 |
| HDL-C | 0.51 | 0.29 | 1.02 | 0.022 |
| Age | 0.4 | 0.22 | 0.75 | <0.001 |
| eGFR | 0.34 | 0.19 | 0.61 | 0.002 |
| BMI | 1.84 | 1.02 | 3.31 | 0.019 |
| UACR | 5.62 | 3.22 | 9.81 | <0.001 |

*BMI, body mass index; FBG, fasting blood glucose; SCr, serum creatinine; eGFR, estimation of glomerular filtration rate; UACR, urine albumin creatinine ratio; TC, total cholesterol; HDL-C, high-density lipoproteins cholesterol. Showing variables with p-values <0.05.*
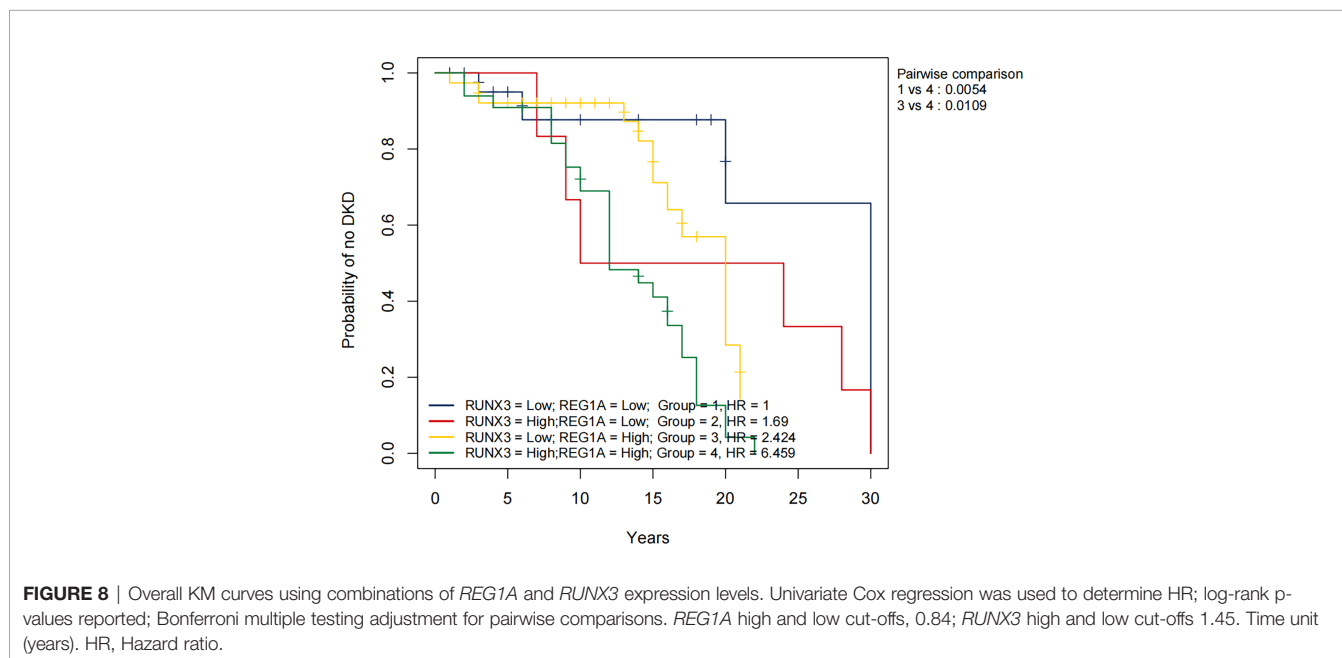
Inhibition of *Runx3* expression ameliorates vascular endothelial dysfunction in DM mice (33). In diabetic patients, *RUNX3* was similarly found to be involved in diabetic endothelial progenitor cell dysfunction (34). High expression of *RUNX3* may exacerbate diabetic vascular endotheliopathy and ultimately lead to DKD. The REG family proteins are structurally similar to each other, and classified as calcium-dependent lectins (35). The mainstream view is that *REG1A* is an indicator of islet β-cell apoptosis and its elevation indicates a decline in islet function (36). In contrast, Okamoto et al. found that *REG1A* stimulates islet regeneration by inducing cell proliferation (37). Sobajima et al. reported significantly increased urinary *REG1A* levels in patients with DKD (38). Li et al. found that REG1A levels are significantly upregulated in the serum of patients with DKD, and possibly associated with renal injury as well (39). The above suggests that REG1A level is associated with impairment of islet function and DKD renal function. The above suggests that REG1A level is associated with islet and DKD renal injury. Our study also found

that serum *REG1A* was significantly and positively associated with renal and islet impairment, while serum *RUNX3* was only associated with renal impairment. Our study also found that serum *REG1A* was significantly elevated in DM, whereas *RUNX3* was unchanged (**Supplementary Figure 4**). We therefore speculate that *REG1A* causes diabetes by destroying islet cells, whereas *RUNX3* causes DKD by directly destroying vascular endothelial cells.

To assess the risk of DKD development, we plotted KM curves for *REG1A*, *RUNX3* and the clinical characteristics. The results showed that people with high *REG1A* and *RUNX3* expression are at an increased risk of DKD after approximately 12 and 8 years of DM, respectively. Our findings are consistent with the results of most previous studies (40–44), which revealed increased TC, FBG, SCr, BMI, and UACR levels as risk factors for DKD, and increased eGFR and HDL-C levels are protective factors for DKD. Contrary to previous findings (45), aging was found to be a protective factor against DKD in this study. Afkarian et al. reported that youth-onset type 2 DM patients may have a higher risk of DKD (46). **Figure 8** shows the KM curves for different expression levels of *REG1A* and *RUNX3*. Accordingly, the risk of DKD was found to be highest when expression levels of both *REG1A* and *RUNX3* are high compared to other groups. Compared to patients with low expression of both *REG1A* and *RUNX3*, DM patients with high expression of both genes are at a rapidly increasing risk of developing DKD after 7-8 years of DM. Hence, *REG1A* and *RUNX3* are potential biomarkers for predicting the risk of developing DKD.

## Limitation

In this study, when comparing the diagnostic efficacy of DEGs in chronic kidney disease (GSE72326), insufficient sample size may have led to unreliable results. We plan to continue collecting blood from patients with chronic kidney disease in subsequent



**FIGURE 8 |** Overall KM curves using combinations of *REG1A* and *RUNX3* expression levels. Univariate Cox regression was used to determine HR; log-rank p-values reported; Bonferroni multiple testing adjustment for pairwise comparisons. *REG1A* high and low cut-offs, 0.84; *RUNX3* high and low cut-offs 1.45. Time unit (years). HR, Hazard ratio.

studies to determine the specificity of *RUNX3* and *REG1A* in the diagnosis of DKD. We have so far only found elevated *RUNX3* and *REG1A* at the transcriptional level in the kidney and subsequent collection of kidney tissue from DKD patients for immunohistochemistry and protein immunoblotting is required.

## Conclusion

*REG1A* and *RUNX3* were found to have high diagnostic efficacy for DK, which was proven through external validation. *REG1A* and *RUNX3* levels were positively and negatively correlated with UACR and eGFR levels, respectively. Thus, the transcription levels of *REG1A* and *RUNX3* in blood samples have potential to predict DKD risk.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Shenzhen people's hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

XW, HW and GY contributed equally to this work. XW, GY, HW and JX contributed to the acquisition of data, analysis and interpretation of data. SY and XW drafted the work or revised it critically for important intellectual content. LX, LZ, TL, and XZ, and LK analyzed the data and revised the article critically for important intellectual content. ZL and SY contributed to the conception and the study design. All authors gave their approval of the version to be published. All authors contributed to the article and approved the submitted version.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2022.935796/full#supplementary-material

## REFERENCES

1. Molitch ME, Adler AI, Flyvbjerg A, Nelson RG, So WY, Wanner C, et al. Diabetic Kidney Disease: A Clinical Update From Kidney Disease: Improving Global Outcomes. *Kidney Int* (2015) 87(1):20–30. doi: 10.1038/ki.2014.128
2. Chan L, Nadkarni GN, Fleming F, McCullough JR, Connolly P, Mosoyan G, et al. Derivation and Validation of a Machine Learning Risk Score Using Biomarker and Electronic Patient Data to Predict Progression of Diabetic Kidney Disease. *Diabetologia* (2021) 64(7):1504–15. doi: 10.1007/s00125-021-05444-0
3. American Diabetes Association.Standards of Medical Care in Diabetes-2016: Summary of Revisions. *Diabetes Care* (2016) 39 Suppl 1:S4–5. doi: 10.2337/dc16-S003
4. Tuttle KR, Bakris GL, Bilous RW, Chiang JL, de Boer IH, Goldstein-Fuchs J, et al. Diabetic Kidney Disease: A Report From an ADA Consensus Conference. *Diabetes Care* (2014) 37(10):2864–83. doi: 10.2337/dc14-1296
5. Hellemons ME, Kerschbaum J, Bakker SJ, Neuwirt H, Mayer B, Mayer G, et al. Validity of Biomarkers Predicting Onset or Progression of Nephropathy in Patients With Type 2 Diabetes: A Systematic Review. *Diabetes Med* (2012) 29 (5):567–77. doi: 10.1111/j.1464-5491.2011.03437.x
6. Coca SG, Nadkarni GN, Huang Y, Moledina DG, Rao V, Zhang J, et al. Plasma Biomarkers and Kidney Function Decline in Early and Established Diabetic Kidney Disease. *J Am Soc Nephrol* (2017) 28(9):2786–93. doi: 10.1681/ASN.2016101101
7. Gohda T, Niewczas MA, Ficociello LH, Walker WH, Skupien J, Rosetti F, et al. Circulating TNF Receptors 1 and 2 Predict Stage 3 CKD in Type 1 Diabetes. *J Am Soc Nephrol* (2012) 23(3):516–24. doi: 10.1681/ASN.2011060628
8. Kar D, Gillies C, Nath M, Khunti K, Davies MJ, Seidu S, et al. Association of Smoking and Cardiometabolic Parameters With Albuminuria in People With Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis. *Acta Diabetol* (2019) 56(8):839–50. doi: 10.1007/s00592-019-01293-x

9. Mayer G, Heerspink HJ, Aschauer C, Heinzel A, Heinze G, Kainz A, et al. Systems Biology-Derived Biomarkers to Predict Progression of Renal Function Decline in Type 2 Diabetes. *Diabetes Care* (2017) 40(3):391–7. doi: 10.2337/dc16-2202
10. Looker HC, Colombo M, Hess S, Brosnan MJ, Farran B, Dalton RN, et al. Biomarkers of Rapid Chronic Kidney Disease Progression in Type 2 Diabetes. *Kidney Int* (2015) 88(4):888–96. doi: 10.1038/ki.2015.199
11. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K., et al. Transcriptome Analysis of Human Diabetic Kidney Disease. *Diabetes* (2011) 60(9):2354–69. doi: 10.2337/db10-1181
12. Sur S, Nguyen M, Boada P, Sigdel TK, Sollinger H, Sarwal MM. FcER1: A Novel Molecule Implicated in the Progression of Human Diabetic Kidney Disease. *Front Immunol* (2021) 12:769972. doi: 10.3389/fimmu.2021.769972
13. Jourde-Chiche N, Whalen E, Gondouin B, Speake C, Gersuk V, Dussol B, et al. Modular Transcriptional Repertoire Analyses Identify a Blood Neutrophil Signature as a Candidate Biomarker for Lupus Nephritis. *Rheumatol (Oxford)* (2017) 56(3):477–87. doi: 10.1093/rheumatology/kew439
14. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, et al. NCBI GEO: Mining Millions of Expression Profiles—Database and Tools. *Nucleic Acids Res* (2005) 33(suppl_1):D562–6. doi: 10.1093/nar/gki022
15. Patwardhan N, Siever E, Spainhour S, Perl. O'Reilly Media, Inc. (2002).
16. Smyth GK. *Limma: Linear Models for Microarray Data, in Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* (The Walter and Eliza Hall Institute of Medical Research, Melbourne Australia: Springer) (2005) p. 397–420.
17. Wickham H. Ggplot2. *Wiley Interdisc reviews: Comput Stat* (2011) 3 (2):180–5. doi: 10.1002/wics.147
18. Ranstam J, Cook J. LASSO Regression. *J Br Surg* (2018) 105(10):1348–8. doi: 10.1002/bjs.10895
19. Hastie T, Qian J, Tay K. *Package 'Glmnet'. CRAN R Repositary.* (2021).
20. Park SH, Goo JM, Jo CH. Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean J Radiol* (2004) 5(1):11–8. doi: 10.3348/kjr.2004.5.1.11

21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC. *Package 'Proc'*, Vol. 09 (University of Basel, Switzerland: Package 'pROC). (2021). p. 34, 2012-09-10.

22. Kramer AA, Zimmerman JE. Assessing the Calibration of Mortality Benchmarks in Critical Care: The Hosmer-Lemeshow Test Revisited. *Crit Care Med* (2007) 35(9):2052–6. doi: 10.1097/01.CCM.0000275267.64078.B0

23. Harrell FEJr., Harrell MFEJr., Hmisc D. *Package 'rms*. (America: Vanderbilt University) (2017). p. 229.

24. Tollefson M. *Graphics with the ggplot2 Package: An Introduction. In: Visualizing Data in R 4*. (Berkeley, CA: Apress) (2021).

25. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A Practical Guide to Understanding Kaplan-Meier Curves. *Otolaryngology— Head Neck Surg* (2010) 143(3):331–6. doi: 10.1016/j.otohns.2010.05.007

26. Bjornstad P, Pyle L, Cherney DZI, Johnson RJ, Sippl R, Wong R, et al. Plasma Biomarkers Improve Prediction of Diabetic Kidney Disease in Adults With Type 1 Diabetes Over a 12-Year Follow-Up: CACTI Study. *Nephrol Dial Transplant* (2018) 33(7):1189–96. doi: 10.1093/ndt/gfx255

27. Ng DP, Salim A, Liu Y, Zou L, Xu FG, Huang S, et al. A Metabolomic Study of Low Estimated GFR in non-Proteinuric Type 2 Diabetes Mellitus. *Diabetologia* (2012) 55(2):499–508. doi: 10.1007/s00125-011-2339-6

28. Morita M, Hanai K, Uchigata Y. Urinary Type IV Collagen as a Predictor for the Incidence of Microalbuminuria in Young Patients With Type 1 Diabetes. *Diabetes Med* (2014) 31(2):213–8. doi: 10.1111/dme.12317

29. Ito Y, Bae SC, Chuang LS. The RUNX Family: Developmental Regulators in Cancer. *Nat Rev Cancer* (2015) 15(2):81–95. doi: 10.1038/nrc3877

30. Chen F, Liu X, Bai J, Pei D, Zheng J. The Emerging Role of RUNX3 in Cancer Metastasis (Review). *Oncol Rep* (2016) 35(3):1227–36. doi: 10.3892/or.2015.4515

31. Smyth LJ, Kilner J, Nair V, Liu H, Brennan E, Kerr K, et al. Assessment of Differentially Methylated Loci in Individuals With End-Stage Kidney Disease Attributed to Diabetic Kidney Disease: An Exploratory Study. *Clin Epigenet* (2021) 13(1):99. doi: 10.1186/s13148-021-01081-x

32. Mou X, Zhou DY, Zhou D, Liu K, Chen LJ, Liu WH, et al. A Bioinformatics and Network Pharmacology Approach to the Mechanisms of Action of Shenxiao Decoction for the Treatment of Diabetic Nephropathy. *Phytomedicine* (2020) 69:153192. doi: 10.1016/j.phymed.2020.153192

33. Jin Q, Lin L, Zhao T, Yao X, Teng Y, Zhang D, et al. Overexpression of E3 Ubiquitin Ligase Cbl Attenuates Endothelial Dysfunction in Diabetes Mellitus by Inhibiting the JAK2/STAT4 Signaling and Runx3-Mediated H3k4me3. *J Transl Med* (2021) 19(1):469. doi: 10.1186/s12967-021-03069-w

34. Meng S, Cao J, Zhang X, Fan Y, Fang L, Wang C, et al. Downregulation of microRNA-130a Contributes to Endothelial Progenitor Cell Dysfunction in Diabetic Patients *via* its Target Runx3. *PloS One* (2013) 8(7):e68611. doi: 10.1371/journal.pone.0068611

35. Zhu X, Dong B, Reding T, Peng Y, Lin H, Zhi M, et al. Association of Serum PSP/REG Iα With Renal Function in Pregnant Women. *BioMed Res Int* (2019) 2019:6970890. doi: 10.1155/2019/6970890

36. Bacon S, Kyithar MP, Schmid J, Rizvi SR, Bonner C, Graf R, et al. Serum Levels of Pancreatic Stone Protein (PSP)/reg1A as an Indicator of Beta-Cell Apoptosis Suggest an Increased Apoptosis Rate in Hepatocyte Nuclear Factor 1 Alpha (HNF1A-MODY) Carriers From the Third Decade of Life Onward. *BMC Endocr Disord* (2012) 12:13. doi: 10.1186/1472-6823-12-13

37. Okamoto H, Takasawa S. Recent Advances in the Okamoto Model: The CD38-Cyclic ADP-Ribose Signal System and the Regenerating Gene Protein (Reg)-Reg Receptor System in Beta-Cells. *Diabetes* (2002) 51 Suppl 3:S462–73. doi: 10.2337/diabetes.51.2007.s462

38. Sobajima H, Niwa T, Shikano M, Naruse S, Kitagawa M, Nakae Y, et al. Urinary Excretion of Pancreatic Stone Protein in Diabetic Nephropathy. *Intern Med* (1998) 37(6):500–3. doi: 10.2169/internalmedicine.37.500

39. Li L, Jia D, Graf R, Yang J. Elevated Serum Level of Pancreatic Stone Protein/Regenerating Protein (PSP/reg) is Observed in Diabetic Kidney Disease. *Oncotarget* (2017) 8(24):38145–51. doi: 10.18632/oncotarget.16369

40. Macisaac RJ, Ekinci EI, Jerums G. Markers of and Risk Factors for the Development and Progression of Diabetic Kidney Disease. *Am J Kidney Dis* (2014) 63(2 Suppl 2):S39–62. doi: 10.1053/j.ajkd.2013.10.048

41. Huang WH, Chen CY, Lin JL, Lin-Tan DT, Hsu CW, Yen TH, et al. High Body Mass Index Reduces Glomerular Filtration Rate Decline in Type II Diabetes Mellitus Patients With Stage 3 or 4 Chronic Kidney Disease. *Med (Baltimore)* (2014) 93(7):e41. doi: 10.1097/MD.0000000000000041

42. Altemtam N, Russell J, El Nahas M. A Study of the Natural History of Diabetic Kidney Disease (DKD). *Nephrol Dial Transplant* (2012) 27(5):1847–54. doi: 10.1093/ndt/gfr561

43. Zoppini G, Targher G, Chonchol M, Ortalda V, Negri C, Stoico V, et al. Predictors of Estimated GFR Decline in Patients With Type 2 Diabetes and Preserved Kidney Function. *Clin J Am Soc Nephrol* (2012) 7(3):401–8. doi: 10.2215/CJN.07650711

44. Chang YH, et al. High-Density Lipoprotein Cholesterol and the Risk of Nephropathy in Type 2 Diabetic Patients. *Nutr Metab Cardiovasc Dis* (2013) 23(8):751–7. doi: 10.1016/j.numecd.2012.05.005

45. Elley CR, Robinson T, Moyes SA, Kenealy T, Collins J, Robinson E, et al. Derivation and Validation of a Renal Risk Score for People With Type 2 Diabetes. *Diabetes Care* (2013) 36(10):3113–20. doi: 10.2337/dc13-0190

46. Afkarian M. Diabetic Kidney Disease in Children and Adolescents. *Pediatr Nephrol* (2015) 30(1):65–74. doi: 10.1007/s00467-014-2796-5