Bioinformatics prediction and experimental verification of key biomarkers for diabetic kidney disease: based on transcriptome sequencing

Jing Zhao [Equal first author, 1, 2], **Kaiying He** [Equal first author, 1, 2], **Hongxuan Du** [1, 2], **Guohua Wei** [2], **Yuejia Wen** [1, 2], **Jiaqi Wang** [1], **Xiaochun Zhou** [Corresp., 2], **Jianqin Wang** [Corresp. 2]

[1] Lanzhou University, Lanzhou, China

[2] Lanzhou University Second Hospital, Lanzhou, China

Corresponding Authors: Xiaochun Zhou, Jianqin Wang
Email address: ery_zhouxc@lzu.edu.cn, ery_wangjqery@lzu.edu.cn

Background Diabetic kidney disease (DKD) is the leading cause of death in people with type 2 diabetes mellitus (T2DM). The main objective of this study is to find the potential biomarkers for DKD. **Materials and Methods**Two datasets (GSE86300 and GSE184836) retrieved from Gene Expression Omnibus (GEO) database were used, combined with our RNA sequencing (RNA-seq) results of DKD mice (C57 BLKS-26W db/db) and non-diabetic (db/m) mice for further analysis. After processing the expression matrix of the 3 sets of data using R software "Limma", differential expression analysis was performed. The significantly differentially expressed genes（DEGs）(|logFC| > 1, *p*-value < 0.05) were visualized by heatmaps and volcano plots respectively. Next, the co-expression genes expressed in the 3 groups of DEGs were obtained by constructing a Venn diagram. In addition, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were further analyzed the related functions and enrichment pathways of these co-expression genes. Then, qRT-PCR was used to verify the expression levels of co-expression genes in the kidney of DKD and control mice. Finally, protein-protein interaction network (PPI), GO, KEGG analysis and Pearson correlation test were performed on the experimentally validated genes, in order to clarify the possible mechanism of them in DKD disease. **Results** Our RNA-seq results identified a total of 125 DEGs, including 59 up-regulated and 66 down-regulated DEGs. At the same time, 183 up-regulated and 153 down-regulated DEGs were obtained in GEO database GSE86300, and 76 up-regulated and 117 down-regulated DEGs were obtained in GSE184836. Venn diagram showed that 13 co-expression DEGs among the 3 groups of DEGs. GO analysis showed that biological processes (BP) were mainly enriched inresponse to stilbenoid, response to fatty acid, response to nutrient, positive regulation of macrophage derived foam cell differentiation, triglyceride metabolic process. KEGG pathway analysis showed

that the 3 major enriched pathways were cholesterol metabolism, drug metabolism - cytochrome P450, PPAR signaling pathway. After qRT-PCR validation, we obtained 11 genes that were significant differentially expressed in the kidney tissues of DKD mice compared with control mice. (The mRNA expression levels of Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Cyp7b1, Akr1c14 and Apoh were declined, whereas Abcc4 and Gsta2 were elevated.) **Conclusion** Our study, based on RNA-seq results, GEO databases and qRT-PCR, identified 11 significant dysregulated DEGs, which play an important role in lipid metabolism and provide novel targets for diagnosis and treatment of DKD.

Bioinformatics prediction and experimental verification of key biomarkers for diabetic kidney disease: based on transcriptome sequencing

Jing Zhao[a,b*], Kaiying He[a,b*], Hongxuan Du[a,b], Guohua Wei[b] , Yuejia Wen[a,b] , Jiaqi Wang[a], Xiaochun Zhou[b#], Jianqin Wang[b#]

a.Lanzhou University, Lanzhou, Gansu, China

b.Department of Nephrology, Lanzhou University Second Hospital, Lanzhou, Gansu, China

Short Title: Key biomarkers for diabetic kidney disease from transcriptome sequencing and bioinformatics analysis

#Corresponding Author:

Xiaochun Zhou

Department of Nephrology

Lanzhou University Second Hospital

No. 82, Cuiyingmen

Lanzhou,Gansu, 730030,China

Tel:15214057999

E-mail: ery_zhouxc@lzu.edu.cn

Jianqin Wang

Department of Nephrology

Lanzhou University Second Hospital

No. 82, Cuiyingmen

Lanzhou,Gansu, 730030,China

Tel:13919038189

E-mail: ery_wangjqery@lzu.edu.cn

#*Equal study contribution

Number of Tables:2

Number of Figures:6

Number of supplementary Tables:2

Key words: DKD; RNA-seq; bioinformatics analysis; differentially expressed genes (DEGs); biomarkers.

Abstract

Background Diabetic kidney disease (DKD) is the leading cause of death in people with type 2 diabetes mellitus (T2DM). The main objective of this study is to find the potential biomarkers for DKD.

Materials and Methods Two datasets (GSE86300 and GSE184836) retrieved from Gene Expression Omnibus (GEO) database were used, combined with our RNA sequencing (RNA-seq) results of DKD mice (C57 BLKS-26W db/db) and non-diabetic (db/m) mice for further analysis. After processing the expression matrix of the 3 sets of data using R software "Limma", differential expression analysis was performed. The significantly differentially expressed genes（DEGs）(|logFC| > 1, p-value < 0.05) were visualized by heatmaps and volcano plots respectively. Next, the co-expression genes expressed in the 3 groups of DEGs were obtained by constructing a Venn diagram. In addition, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were further analyzed the related functions and enrichment pathways of these co-expression genes. Then, qRT-PCR was used to verify the expression levels of co-expression genes in the kidney of DKD and control mice. Finally, protein-protein interaction network (PPI), GO, KEGG analysis and Pearson correlation test were performed on the experimentally validated genes, in order to clarify the possible mechanism of them in DKD disease.

Results Our RNA-seq results identified a total of 125 DEGs, including 59 up-regulated and 66 down-regulated DEGs. At the same time, 183 up-regulated and 153 down-regulated DEGs were obtained in GEO database GSE86300, and 76 up-regulated and 117 down-regulated DEGs were obtained in GSE184836. Venn diagram showed that 13 co-expression DEGs among the 3 groups of DEGs. GO analysis showed that biological processes (BP) were mainly enriched inresponse to stilbenoid, response to fatty acid, response to nutrient, positive regulation of macrophage derived foam cell differentiation, triglyceride metabolic process. KEGG pathway analysis showed that the 3 major enriched pathways were cholesterol metabolism, drug metabolism - cytochrome P450, PPAR signaling pathway. After qRT-PCR validation, we obtained 11 genes that were significant differentially expressed in the kidney tissues of DKD mice compared with control mice. (The mRNA expression levels of Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Cyp7b1, Akr1c14 and Apoh were declined, whereas Abcc4 and Gsta2 were elevated.)

Conclusion Our study, based on RNA-seq results, GEO databases and qRT-PCR, identified 11 significant dysregulated DEGs, which play an important role in lipid metabolism and provide novel targets for diagnosis and treatment of DKD.

1.Introduction

Diabetic mellitus (DM) is a chronic metabolic disease that seriously affects public health. According to the World Health Organization (WHO), about 629 million people will suffer from T2DM by 2045, and the complications may affect various organs throughout the body and have a high mortality and disability rate (Tanase et al. 2020). Diabetic kidney disease (DKD) is a common chronic complication of DM, which has gradually been the leading cause of death in patients with DM and end-stage renal disease (ESRD)(Yang et al. 2020). The main features of progression of DKD are hypertension, increased protein in urine, and decreased estimated glomerular filtration rate (eGFR). Early pathological changes mainly include thickening of the glomerular and tubular basement membrane, which gradually develops into glomerular extracellular matrix (ECM) accumulation and tubular interstitial fibrosis as the disease progresses, eventually causing irreversible

damage to the renal structure. Most experts believe that the pathogenesis of DKD is mainly due to the interaction of renin-angiotensin-aldosterone system (RAAS), advanced glycation end products (AGEs), transforming growth factor-β1 (TGF-β1), protein kinase C (PKC), mitogen-activated protein kinases (MAPKs) and reactive oxygen species (ROS), which affect renal function through inflammation and oxidative stress(Samsu 2021). Although various treatments have been used to improve metabolism, hemodynamic disturbances, and fibrosis, the mortality rate of patients with DKD remains high. Therefore, the key to improving the quality of life and survival rate is to find more useful biomarkers for diagnosis of DKD and to develop new strategies and prevention of deterioration of renal function.

In recent years, many biomarker molecules have been found to be associated with changes in renal structure and function in DKD patients, such as urine markers, serum/plasma markers, etc(Colhoun & Marcovecchio 2018). With the development of the new generation of high-throughput sequencing technology and bioinformatics techniques, the ability of humans to understand diseases from the root has greatly improved, and more and more disease-related risk genes have been discovered. This promises revolutionary advances in disease diagnosis and treatment in the future (Rego & Snyder 2019). Many microarray-based studies have shown that non-coding RNAs, mRNAs and the protein it encodes play important roles in the pathogenesis of DKD. They influence disease occurrence, progression, and prognosis through their interactions and regulation of signaling pathways. For example, studies have shown that leucine rich-α-2 glycoprotein 1 (LRG1) expression is increased in kidneys of diabetic patients and mice, it can be used as a biomarker for diagnosis of DKD(Hong et al. 2019). The expression of SH3YL1 in the serum of DKD patients, in kidney tissues of db/db mice, and in podocytes, is higher than control group, so it also has the potential as a biomarker for the diagnosis of DKD (Choi et al. 2021). Some nephrologists have used bioinformatics methods to analyze the datasets uploaded by other authors in GEO database, and screened out candidate genes in DEGs through functional analysis and protein interaction network construction, which is considered as a possible biomarker of DKD(Gao et al. 2021). The pathogenesis of DKD is complex, which is related to oxidative stress, inflammation, autophagy, apoptosis, and other mechanisms. However, its pathogenesis still needs further study. Although many mRNAs and proteins it encodes are currently thought to play a role in DKD, there are few consistent results across studies. We need to study by using the method of biological technology, such as through such as single cell sequencing, high throughput sequencing, multi-omics combined analysis to obtain more data analysis, or we share our own raw data in the database so that more researchers can study it. Then, some basic experiments are added to verify the mechanism of genes action in diseases, such as qRT-PCR and Western Blot, so that the research results on the mechanism of genes in DKD can have higher authenticity and reliability.

Unlike most bioinformatics studies, we added our own sequencing data, and performed a comprehensive analysis with data similar to our sequencing uploaded by other authors in the GEO database. The candidate gene was verified by qRT-PCR again, increasing the accuracy of the study. Our analysis was conducted following the procedure presented in Fig 1. Finally, 13 co-expression genes were determined by the combined analysis of the 3 datasets, qRT-PCR verified that 11 of them had the same expression trend as the above sequencing results and analysis showed several of these genes may affect the occurrence of disease through lipid metabolism and PPAR Signaling pathway.

2.Materials and Methods

2.1 Data source

2.1.1 Transcriptome sequencing of kidney tissues from DKD and Control mice

1) Animal model

Twelve C57 BLKS-db/db mice and twelve db/m mice（6-week-old,male）were selected as the model group (body weight 34.35±2.52g),normal control group(body weight 18.94±1.44g) respectively. They were purchased from Nanjing Institute of Model Zoology and reared in the barrier system of Animal Experimental Research Centre of Lanzhou University Second Hospital. The feeding temperature was (20±2) ℃, the humidity was 40%-70%, the light alternated between light and dark every 12 hours, the ordinary Specific Pathogen Free (SPF) food was fed and drank water freely. Starting from the 8th week, the body weight, blood sugar (Sinocare Inc), and 16-hour urine microalbumin (Enzyme Linked Immunosorbent Assay (ELISA) kit of mlbio) of the mice were measured every 4 weeks, the blood was collected from the tail vein. We got blood from the heart at the 8th and 32nd weeks (six mice were randomly selected for blood collection from heart at week 8, and the other six were fed to 32 weeks for blood collection in heart), stored at 4°C for 1 hour, and centrifuged for 10 minutes (3000 rpm, 4℃) to obtain the serum. Then the serum creatinine (Scr), blood urea nitrogen (BUN), total cholesterol (TC), triglycerides (TG), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C) were measured by ELISA kit of Nanjing Jiancheng Bioengineering Institute. We fed the mice with a normal diet for 26 weeks, then the mice were euthanized by intraperitoneal injection of 10% chloral hydrate and the kidney tissues were collected. 50 mg of the right kidney was divided into cryogenic vials, frozen with liquid nitrogen and stored in the refrigerator at -80°C. All the experiments involving animals were performed after Lanzhou University Second Hospital Institutional Ethical Committee's approval and under the strict adherence to the National Institutes of Health Guide for laboratory animals' care and use, our ethical review acceptance number is D2019-198.

2) RNA extraction from kidney tissues

Approximately 50 mg of each kidney tissue sample was placed in grinding tubes, and 1 ml of Trizol was added to each tube and ground completely. Then the samples were allowed to stand for 5 minutes at room temperature, and 0.2 ml of chloroform was added to each tube and shaken vigorously for 15 seconds. After the samples stand at room temperature for 3 minutes, the supernatant was centrifuged at high speed (12000 rpm, 4℃) for 15 minutes and transferred to a Rnase-free centrifuge tube. Next, add 0.5 mL isopropyl alcohol, mix gently and let stand at room temperature for 10 minutes, centrifuge for 10 minutes (12000 rpm, 4℃), wash the precipitate with 1 mL 75% ethanol, centrifuge for 5 minutes (7500 rpm, 4℃), remove the liquid, dry at room temperature for 10 minutes. 150uL diethyl pyrocarbonate (DEPC) H_2O was added and gently mix. At last, the precipitate was left at 55℃ for 10 minutes to dissolve and stored temporarily at -80℃ for further sequencing.

3) RNA-sequencing（RNA-seq）

In this study, we need to consider and deal with the difference expression caused by the biological variability, by far the most commonly used and most effective way is to set up biological replic validates. All biological repeat samples under the same conditions were extracted and built with the same person and batch, sequenced with same Run and Lane, and conducted a detailed analysis of the abnormal samples. Meanwhile, we set the Power value as 0.9, α value as 0.05 and then used the R language RNAseqPower package to calculate the sample size was 2.44, which could achieve 1.5 times of change, proving that the number of mice in each group selected as 3 is valid in our experiment.

Illumina's high-throughput sequencing platform was used to perform transcriptome sequencing on kidney tissue samples from DKD mice and Control mice. Clean data were obtained by filtering data from the Illumina high-throughput sequencing platform sequenced with the indicated reference genome. Next, the sequencing data were

obtained through sequence structure analysis, library quality assessment and comparing with the reference genome. Finally, the number of mapped reads and transcript length were normalized in the sample. Fragments Per Kilobase Million (FPKM) was used as an indicator to measure gene expression levels. The expression matrix of FPKM can be obtained by computational formula shown as:FPKM=(cDNA Fragments\over(Mapped Fragments (Millions)*Transcript Length(kb))).

2.1.2 GEO database

GEO database was built by the National Center for Biotechnology Information (NCBI) and is a gene expression database and online genome resource that collects high-throughput gene expression data uploaded from research institutes around the world. Diabetic nephropathy or diabetic kidney disease was entered as search objects. Gene expression microarray datasets GSE86300 and GSE184836 were selected and downloaded. The criteria for selecting the datasets were as follows: DKD and Control mice models with detailed gene expression information. The GSE86300 dataset, based on the GPL7546 platform, includes 10 DKD kidney tissue samples and 10 Control kidney tissue samples. The GSE184836 dataset, based on the GPL21103 platform, includes 6 DKD kidney tissue samples and 6 Control kidney tissue samples. The detailed information on these microarray datasets is listed in Supplemental Table 1.

2.2 Data processing methods

2.2.1 Differential expression analysis

All differential analyses of the 3 datasets were performed using the Limma packages in R/Bioconductor software, we uploaded the script for the Limma package to the supplement file (supplement file). Genes in DKD renal tissue were up- or down-regulated compared with control groups, *p*-value less than 0.05 and $|\log_2$ fold change (FC)| greater than 1 were considered statistically significant in differential analysis of the 3 datasets, logFC＞1 was regarded as up-regulated genes and logFC＜-1 was down-regulated.

2.2.2 Volcano plots and heatmaps

Significant DEGs (|logFC| > 1, *p*-value < 0.05) were visualized using heatmaps and volcano curves. Both visualizations were completed by http://www.bioinformatics.com.cn, an online platform for data analysis and visualization. We drew heatmaps by importing FPKM values of the two groups to the website, and using the same method, we imported gene names, logFC values, and *P*-value to map the volcano.

2.2.3 Venn diagram

Determine the common genes of 3 datasets by creating a Venn diagram. The Venn diagram is completed on the "Weishengxin" website. We imported three sets of datasets to the Venn diagram tool on the website, draw Venn diagrams, and obtain the intersection.

2.2.4 Functional enrichment analysis

Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed for DEGs. GO Annotation included analysis of biological processes (BP), cell components (CC), and molecular functions (MF). KEGG is an online database for pathway analysis of a large amount of genetic information. GO functional annotation analysis and KEGG pathway enrichment analysis were performed in the website "Weishengxin", import our gene name and logFC into the website for enrichment analysis. The significant enrichment threshold was set as *p*-value<0.05, ten functions and pathways with the lowest *p*-values were selected as the top 10.

2.2.5 Quantitative real-time polymerase chain reaction（qRT-PCR）

Total RNA was extracted from fresh mice kidneys of DKD and control groups using the TRIZOL method. Total

RNA (1 ug) was transcribed into cDNA using the GoScript™ Reverse Transcription System according to the manufacturer's protocol (Promega). qRT-PCR was performed on the ABI7500 system using the GoTaq® qPCR Master Mix (Promega). All data were normalized to β-actin expression. Relative RNA expression was calculated using the $2^{-\Delta\Delta CT}$ method. Detailed information of primers was listed in Supplemental Table 2.

2.2.6 PPI network analysis and Pearson correlation test.

The protein-protein interaction (PPI) network was created from the "STRING" database for target genes and is designed to discover the interaction of target genes and their interactions with other proteins, and Cytoscape was used to edit the graphics. The expression matrix of the 3 datasets was merged after removing batch effect by the Sangerbox tool, and the Pearson correlation test was adopted for evaluation of the interactions between DKD related genes at the mRNA level using R.

2.3 Statistical analysis

All values are expressed as the mean±SEM. Statistical analysis was performed using the statistical package SPSS for Windows Version 7.51 (SPSS, Inc, Chicago, IL, USA) and GraphPad Prism 9. Results were analyzed using Student t test for multiple comparisons. In our qRT-PCR experiment, each sample was added with three holes each time, and three groups of different samples were repeated for verification. Finally, we performed T test on the $2^{-\Delta\Delta CT}$ of the three groups of data to obtain the *P*-value. Statistical significance was detected at the 0.05 level.

3.Results

3.1 General condition and biochemical indexes of mice

From the age of 8 weeks, compared with db/m mice, the water and food intake of db/db mice increased gradually, the body weight increased ($p<0.05$), blood sugar increased ($p<0.05$), urinary protein excretion rate and various biochemical indexes changed significantly ($p<0.05$) (Fig.2).

3.2 Differential expression gene

Based on our RNA-seq results, a total of 125 differentially expressed mRNAs were identified, including 59 up-regulated genes and 66 down-regulated genes. After screening the genes with differential expression in GSE86300, 336 genes with differential expression were determined, including 183 up-regulated genes and 153 down-regulated genes. In GSE184836 DEGs, there were 315 differentially expressed genes, including 76 up-regulated genes and 117 down-regulated genes. The 3 datasets were analyzed according to the criteria $|\log_2 FC| > 1$ and p-value < 0.05. To visualize the differentially expressed genes, we constructed heatmaps (Fig. 3A, 3B, 3C) and volcano curves (Fig. 3D, 3E, 3F) and then created a Venn diagram to obtain 13 co-expression genes among the 3 datasets (Fig. 3G).

3.3 Enrichment analysis

GO functional annotation and KEGG pathway enrichment analysis were performed for the DEGs from the 3 datasets. The GO analysis results of our RNA-seq showed that the BP analysis was mainly enriched in the organic anion transport, fatty acid metabolic process, organic hydroxy compound metabolic process, steroid metabolic process, response to stilbenoid, while CC is mainly concentrated in chylomicron, very-low-density lipoprotein particle, triglyceride-rich plasma lipoprotein particle, plasma lipoprotein particle, lipoprotein particle. The top 5 of MF are monooxygenase activity, oxidoreductase activity, organic anion transmembrane transporter activity, lipid transporter activity, glucuronosyltransferase activity (Fig. 4A). KEGG pathway analysis showed that the top 3 enrichment pathways of the DEGs are Bile secretion, Steroid hormone biosynthesis, drug metabolism - other enzymes (Fig. 4B).

GSE86300 analysis showed that the BP of DEGs was mainly in the negative regulation of blood coagulation,

negative regulation of hemostasis, fibrinolysis (Fig.4C). At the same time, KEGG analysis showed they were mainly in the Cholesterol metabolism, Complement and coagulation cascades, butanoate metabolism, PPAR signaling pathway (Fig.4D). GSE184836 analysis indicated that organic anion transport, organic hydroxy compound metabolic process, steroid metabolic process as the top 3 BP (Fig.4E). Meanwhile, KEGG analysis manifested that metabolism of xenobiotics by cytochrome P450, cholesterol metabolism, butanoate metabolism as the top 3 (Fig.4F). The results of GO and KEGG analysis are sorted by *p*-value.

3.4 Venn diagram

13 overlapping genes, namely, Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Kynu, Cyp7b1, Akr1c14, Apoh, Fmo5, Abcc4, Gsta2 were identified from the 3 datasets by creating a Venn diagram. The expression levels and trends of the above 13 genes in 3 databases were listed in Table1.

3.5 Identification of 13 overlapping genes expression by qRT-PCR

The mRNA levels of the 13 overlapping genes were verified by qRT-PCR in DKD and control mice kidney tissues, meanwhile, the results indicated that the mRNA levels of Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Cyp7b1, Akr1c14, Apoh were down-regulated, while Abcc4 and Gsta2 were up-regulated, these changes were statistically significant ($p < 0.05$). In addition, qRT-PCR results of Fmo5 showed that its expression was up-regulated in DKD kidney tissues, which is contrary to the sequencing results, the expression of Kynu showed no statistical significance in the DKD group compared with the Control ($p > 0.05$) (Fig 5). In conclusion, the mRNA expression levels of these 11 genes verified by qRT-PCR were consistent with the sequencing results (Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Cyp7b1, Akr1c14, Apoh, Abcc4, Gsta2).

3.6 Protein interaction analysis

The “STRING” database was used for PPI analysis. The results of PPI analysis of the above verified 11 genes were similar to those of functional annotation of GO and KEGG Pathway Enrichment Analysis (Fig.6A-C). These genes apparently showed mainly correlation with lipid metabolism, PPAR signaling pathway, and of these 11 genes, 3 had an obvious interaction relationship, namely Cd36, Apoh, and Lpl (Fig.6C). We also got some genes associated with them (Fig.6D). Similarly, the enrichment analysis of the BP of 11 genes revealed that the above 3 genes were all enriched in the triglyceride metabolic process. Additional information of the 11 verified genes related to BP analysis is listed in Table2. In addition, through the Pearson correlation test, there were prominently positive or negative correlations between the DKD related genes at the mRNA levels (Fig. 6E).

4.Discussion

In recent years, the consequences of DKD have gradually attracted people's attention. With the increasing incidence, it brings a heavy medical burden to society. Nowadays, the clinical diagnosis of DKD mainly depends on the protein content in urine, and the treatment aims to reduce the protein content in urine. With the development of sequencing technology, we now realize that genes play a key role in the diagnosis, occurrence, progression and treatment of DKD. Thus, many mRNAs and non-coding RNAs associated with DKD have been discovered and supported by a large number of studies.

In our study, the significant increase of UAER, as well as the changes of Scr and other biochemical indicators, could indicate the decline of renal function of mice. We found 11 genes whose expression levels were significantly altered in the kidney tissues of DKD mice. Aacs, Cpe, Cd36, Slc22a7, Slc1a4, Lpl, Cyp7b1, Akr1c14, Apoh were down-regulated whereas the expression levels of Abcc4 and Gsta2 were upregulated. We found that these genes are homologous genes in humans and mice in the National Center for Biotechnology Information (NCBI) database, and the genes have been studied in human and mice models of different diseases.

Therefore, we have reason to believe that these genes have similar functions in human disease and mice model research, so we can screen important genes on the basis of mice study, and provide some reference data for the study of this disease in human.

In our findings, what attracted our attention was that some genes were involved in lipid metabolism and PPAR signaling pathway. It is well known that glucose and lipid metabolism disorder has become a very common metabolic defect, lipid metabolism disorders are closely associated with renal dysfunction. Researchers have confirmed that dyslipidemia and the increased free fatty acids (FFA) are risk factors for insulin resistance and glucolipid toxicity is an important cause of T2DM(Lytrivi et al. 2020). Lipid toxicity is mainly involved in renal damage related to lipid toxicity through the activation of inflammation, oxidative stress, mitochondrial dysfunction and apoptosis(Opazo-Ríos et al. 2020), and increased lipophagy ameliorates renal damage, lipid deposition, oxidative stress and apoptosis in the kidney(Han et al. 2021). In addition, PPAR signaling pathway has been proved to be significantly related to lipid metabolism by a large number of studies, and its role in DKD has also been widely recognized(Kim et al. 2018; Mao et al. 2021).

The results of PPI showed that five genes were enriched in lipid metabolic process, namely Cd36, Lpl, Apoh, Aacs and Cyp7b1, at the same time, Cd36, Lpl and Apoh were related to each other, which may act together to influence the occurrence and development of DKD. We found that these genes have been extensively reported in the literature in past studies. For example, some studies suggested that Cd36 is involved in fatty acid uptake, apoptosis, angiogenesis, phagocytosis, inflammation and atherosclerosis(Pepino et al. 2014). Prior data indicated that high glucose exacerbates fatty acid uptake and deposition through increased expression of Cd36 via the AKT-PPARγ pathway(Alkhatatbeh et al. 2013; Mitrofanova et al. 2020). Cd36 knockdown prevents renal tubular injury, tubulointerstitial inflammation, and oxidative stress in 16 weeks db/db mice(Hou et al. 2021). On the contrary, the loss of Cd36 in mice leads to phenotypes such as lymphatic drainage, visceral fat, and glucose intolerance, which increases the risk for T2DM(Cifarelli et al. 2021). In our study, compared with the control mice, Cd36 in the DKD group was down-regulated. We speculated the reason might be that our mice had reached the late stage at the age of 32 weeks, so the expression level of DKD was different from that in the early stage. In addition, studies have shown that in C57BL/6J mice, Lpl was mainly expressed intracellular and restricted to the proximal tubule(Nyrén et al. 2019). In the DKD group, renal Lpl mRNA expression was significantly decreased, and it increased the level of triglyceride (TG) in renal tissue(Herman-Edelstein et al. 2014). PPAR is involved in lipid metabolism and it regulates the activity of proteins like LPL(La Paglia et al. 2017). Apoh is involved in a variety of physiological processes, including lipoprotein metabolism, coagulation, and antiphospholipid autoantibody production(Athanasiadis et al. 2013). It is also associated with BMI in diabetics and cardiovascular risk factors and is considered an anti-obesity factor(Hasstedt et al. 2016). In addition to these, Aacs has been widely recognized for its role in cellular response to glucose stimulation, regulation of insulin secretion, and fatty acid metabolism. Previous studies have shown that 80% inhibition of Aacs can partially inhibit glucose-induced insulin release(MacDonald et al. 2007). Besides, the enzyme encoded by Aacs is a ketone body-utilizing ligase with a role in lipid synthesis through the non-oxydative pathway(Haydar et al. 2019), knockdown of Aacs in vivo resulted in the reduction of total blood cholesterol(Hasegawa et al. 2012). Studies have shown that Cyp7b1 is related to the synthesis of bile acid (BA)(Evangelakos et al. 2022), which is the final product of cholesterol catabolism and is related to the regulation of lipid, glucose and energy metabolism, inflammation, detoxification of drug metabolism(Li & Chiang 2014), as well as the development of liver steatosis and inflammation(Evangelakos et al. 2022).

In addition, some genes are not enriched in lipid metabolism and PPAR pathways, and more studies are needed in the future to fill in the gaps in their functional studies. For example, Akr1c14, a metabolic enzyme which has been thought to play a role in the maintenance of normal adipocyte metabolism(Becker et al. 2019). Abcc4 is highly expressed in the kidney of mice, which has been verified that the expression of Mrp4/Abcc4 increased in liver and kidney of obese and diabetic patients(Donepudi et al. 2021). Carboxy peptidase E (Cpe), an enzyme that converts the pro-insulin to insulin, which suggest that Cpe may be related to the occurrence of T2DM(Sabiha et al. 2021), rescuing proinsulinmia caused by reduced CPE may be a new approach to treat early diabetes(Jo et al. 2019). Nuclear factor erythroid 2-related factor 2 (Nrf2) is known to regulate cellular oxidative stress and induce expression of antioxidant genes. Gsta2 is a downstream signaling molecule of Nrf2 and also participates in the antioxidant stress effect of Nrf2(Kim et al. 2015), which was researched in the mice with surgically induced Unilateral ureteral obstruction (UUO). Slc22a7 regulates the absorption, distribution, and excretion of a wide variety of environmental toxins and clinically important drugs as an organic anion transporter (Zhou & You 2007). Slc1a4 is highly expressed in the central nervous system, and relevant studies are still lacking.

The results of PPI suggest that our key genes are not only related to lipid metabolism and PPAR pathways, but also have other related genes. For example, they interact with Angptl4, Apoa1, Apoa2 and Apo b. Angptl4 has been studied as a potent inhibitor of Lpl that regulates cellular uptake of triglycerides and promotes fatty acid oxidation(Zhou et al. 2021). Binod (Aryal et al. 2016) found that the expression of Cd36 increased in Angptl4$^{-/-}$ macrophages. Apo a and Apo b are mainly studied in obesity and cardiovascular diseases, studies suggest that Apoa1, Apo b, and Apo b/Apoa1 ratio have been regarded as the predictors of microvascular and macrovascular complications of diabetes(Moosaie et al. 2020). And these genes are concentrated in PPAR signaling pathway, which is thought to be involved in the regulation of glucose and lipid metabolism, endothelial function and inflammation(Wang et al. 2016). These genes may also have implications in the influence of key genes on disease. Although we have seen less research of these genes on kidney disease, we speculate from our current study that it may play an important role in kidney disease, more experiments will investigate the mechanism in the future.

However, the present study has certain limitation. Firstly, we determined the renal function status of mice by detecting the hematuria biochemical indexes of the mice, but there was no corresponding pathological evidence. In addition, qRT-PCR was simply used to verify their mRNA expression levels in the kidney tissues of DKD mice, while Western Blot was not used to verify their protein expression level, so the target genes should be further verified in experimental studies. The present results were obtained using a bioinformatics screening to identify several DEGs between DKD and control groups, the results showed that several genes not only played an important role in the lipid metabolism of DKD, but also had a close relationship with PPAR signaling pathway. Importantly, the information provided in this study was not limited to the 11 verified genes, but maybe included certain other typical DEGs. The current results provide a worthy resource for future research on DKD.

5.Conclusion

In conclusion, our study, based on RNA-seq results, the GEO database and qRT-PCR, identified 11 significant dysregulated DEGs, which play an important role in lipid metabolism and provide novel targets for diagnosis and treatment of DKD, additional basic and clinical studies are needed to further validate these targets.

References

Alkhatatbeh MJ, Enjeti AK, Acharya S, Thorne RF, and Lincz LF. 2013. The origin of circulating CD36 in type 2 diabetes. *Nutr Diabetes* 3:e59. 10.1038/nutd.2013.1

Aryal B, Rotllan N, Araldi E, Ramírez CM, He S, Chousterman BG, Fenn AM, Wanschel A, Madrigal-Matute J, Warrier N, Martín-Ventura JL, Swirski FK, Suárez Y, and Fernández-Hernando C. 2016. ANGPTL4 deficiency in haematopoietic cells promotes monocyte expansion and atherosclerosis progression. *Nat Commun* 7:12313. 10.1038/ncomms12313

Athanasiadis G, Sabater-Lleal M, Buil A, Souto JC, Borrell M, Lathrop M, Watkins H, Almasy L, Hamsten A, and Soria JM. 2013. Genetic determinants of plasma β_2-glycoprotein I levels: a genome-wide association study in extended pedigrees from Spain. *J Thromb Haemost* 11:521-528. 10.1111/jth.12120

Becker JC, Gérard D, Ginolhac A, Sauter T, and Sinkkonen L. 2019. Identification of genes under dynamic post-transcriptional regulation from time-series epigenomic data. *Epigenomics* 11:619-638. 10.2217/epi-2018-0084

Choi GS, Min HS, Cha JJ, Lee JE, Ghee JY, Yoo JA, Kim KT, Kang YS, Han SY, Bae YS, Lee SR, Yoo JY, Moon SH, Lee SJ, and Cha DR. 2021. SH3YL1 protein as a novel biomarker for diabetic nephropathy in type 2 diabetes mellitus. *Nutr Metab Cardiovasc Dis* 31:498-505. 10.1016/j.numecd.2020.09.024

Cifarelli V, Appak-Baskoy S, Peche VS, Kluzak A, Shew T, Narendran R, Pietka KM, Cella M, Walls CW, Czepielewski R, Ivanov S, Randolph GJ, Augustin HG, and Abumrad NA. 2021. Visceral obesity and insulin resistance associate with CD36 deletion in lymphatic endothelial cells. *Nat Commun* 12:3350. 10.1038/s41467-021-23808-3

Colhoun HM, and Marcovecchio ML. 2018. Biomarkers of diabetic kidney disease. *Diabetologia* 61:996-1011. 10.1007/s00125-018-4567-5

Donepudi AC, Lee Y, Lee JY, Schuetz JD, and Manautou JE. 2021. Multidrug resistance-associated protein 4 (Mrp4) is a novel genetic factor in the pathogenesis of obesity and diabetes. *Faseb j* 35:e21304. 10.1096/fj.202001299RR

Evangelakos I, Kuhl A, Baguhl M, Schlein C, John C, Rohde JK, Heine M, Heeren J, and Worthmann A. 2022. Cold-Induced Lipoprotein Clearance in Cyp7b1-Deficient Mice. *Front Cell Dev Biol* 10:836741. 10.3389/fcell.2022.836741

Gao Z, S A, Li XM, Li XL, and Sui LN. 2021. Identification of Key Candidate Genes and Chemical Perturbagens in Diabetic Kidney Disease Using Integrated Bioinformatics Analysis. *Front Endocrinol (Lausanne)* 12:721202. 10.3389/fendo.2021.721202

Han Y, Xiong S, Zhao H, Yang S, Yang M, Zhu X, Jiang N, Xiong X, Gao P, Wei L, Xiao Y, and Sun L. 2021. Lipophagy deficiency exacerbates ectopic lipid accumulation and tubular cells injury in diabetic nephropathy. *Cell Death Dis* 12:1031. 10.1038/s41419-021-04326-y

Hasegawa S, Noda K, Maeda A, Matsuoka M, Yamasaki M, and Fukui T. 2012. Acetoacetyl-CoA synthetase, a ketone body-utilizing enzyme, is controlled by SREBP-2 and affects serum cholesterol levels. *Mol Genet Metab* 107:553-560. 10.1016/j.ymgme.2012.08.017

Hasstedt SJ, Coon H, Xin Y, Adams TD, and Hunt SC. 2016. APOH interacts with FTO to predispose to healthy thinness. *Hum Genet* 135:201-207. 10.1007/s00439-015-1629-3

Haydar S, Grigorescu F, Vintilă M, Cogne Y, Lautier C, Tutuncu Y, Brun JF, Robine JM, Pugeat M, Normand C, Poucheret P, Gheorghiu ML, Georgescu C, Badiu C, Băculescu N, Renard E, Ylli D, Badiou S, Sutra T, Cristol JP, Mercier J, Gomis R, Macias JM, Litvinov S, Khusnutdinova E, Poiana C, Pasquali R, Lauro D,

Sesti G, Prudente S, Trischitta V, Tsatsoulis A, Abdelhak S, Barakat A, Zenati A, Ylli A, Satman I, Kanninen T, Rinato Y, and Missoni S. 2019. Fine-scale haplotype mapping of MUT, AACS, SLC6A15 and PRKCA genes indicates association with insulin resistance of metabolic syndrome and relationship with branched chain amino acid metabolism or regulation. *PLoS One* 14:e0214122. 10.1371/journal.pone.0214122

Herman-Edelstein M, Scherzer P, Tobar A, Levi M, and Gafter U. 2014. Altered renal lipid metabolism and renal lipid accumulation in human diabetic nephropathy. *J Lipid Res* 55:561-572. 10.1194/jlr.P040501

Hong Q, Zhang L, Fu J, Verghese DA, Chauhan K, Nadkarni GN, Li Z, Ju W, Kretzler M, Cai GY, Chen XM, D'Agati VD, Coca SG, Schlondorff D, He JC, and Lee K. 2019. LRG1 Promotes Diabetic Kidney Disease Progression by Enhancing TGF-β-Induced Angiogenesis. *J Am Soc Nephrol* 30:546-562. 10.1681/asn.2018060599

Hou Y, Wang Q, Han B, Chen Y, Qiao X, and Wang L. 2021. CD36 promotes NLRP3 inflammasome activation via the mtROS pathway in renal tubular epithelial cells of diabetic kidneys. *Cell Death Dis* 12:523. 10.1038/s41419-021-03813-6

Jo S, Lockridge A, and Alejandro EU. 2019. eIF4G1 and carboxypeptidase E axis dysregulation in O-GlcNAc transferase-deficient pancreatic β-cells contributes to hyperproinsulinemia in mice. *J Biol Chem* 294:13040-13050. 10.1074/jbc.RA119.008670

Kim S, Kim SJ, Yoon HE, Chung S, Choi BS, Park CW, and Shin SJ. 2015. Fimasartan, a Novel Angiotensin-Receptor Blocker, Protects against Renal Inflammation and Fibrosis in Mice with Unilateral Ureteral Obstruction: the Possible Role of Nrf2. *Int J Med Sci* 12:891-904. 10.7150/ijms.13187

Kim Y, Lim JH, Kim MY, Kim EN, Yoon HE, Shin SJ, Choi BS, Kim YS, Chang YS, and Park CW. 2018. The Adiponectin Receptor Agonist AdipoRon Ameliorates Diabetic Nephropathy in a Model of Type 2 Diabetes. *J Am Soc Nephrol* 29:1108-1127. 10.1681/asn.2017060627

La Paglia L, Listì A, Caruso S, Amodeo V, Passiglia F, Bazan V, and Fanale D. 2017. Potential Role of ANGPTL4 in the Cross Talk between Metabolism and Cancer through PPAR Signaling Pathway. *PPAR Res* 2017:8187235. 10.1155/2017/8187235

Li T, and Chiang JY. 2014. Bile acid signaling in metabolic disease and drug therapy. *Pharmacol Rev* 66:948-983. 10.1124/pr.113.008201

Lytrivi M, Castell AL, Poitout V, and Cnop M. 2020. Recent Insights Into Mechanisms of β-Cell Lipo- and Glucolipotoxicity in Type 2 Diabetes. *J Mol Biol* 432:1514-1534. 10.1016/j.jmb.2019.09.016

MacDonald MJ, Smith AD, 3rd, Hasan NM, Sabat G, and Fahien LA. 2007. Feasibility of pathways for transfer of acyl groups from mitochondria to the cytosol to form short chain acyl-CoAs in the pancreatic beta cell. *J Biol Chem* 282:30596-30606. 10.1074/jbc.M702732200

Mao Z, Feng M, Li Z, Zhou M, Xu L, Pan K, Wang S, Su W, and Zhang W. 2021. ETV5 Regulates Hepatic Fatty Acid Metabolism Through PPAR Signaling Pathway. *Diabetes* 70:214-226. 10.2337/db20-0619

Mitrofanova A, Fontanella AM, Merscher S, and Fornoni A. 2020. Lipid deposition and metaflammation in diabetic kidney disease. *Curr Opin Pharmacol* 55:60-72. 10.1016/j.coph.2020.09.004

Moosaie F, Firouzabadi FD, Abouhamzeh K, Esteghamati S, Meysamie A, Rabizadeh S, Nakhjavani M, and Esteghamati A. 2020. Lp(a) and Apo-lipoproteins as predictors for micro- and macrovascular complications of diabetes: A case-cohort study. *Nutr Metab Cardiovasc Dis* 30:1723-1731. 10.1016/j.numecd.2020.05.011

Nyrén R, Makoveichuk E, Malla S, Kersten S, Nilsson SK, Ericsson M, and Olivecrona G. 2019. Lipoprotein lipase in mouse kidney: effects of nutritional status and high-fat diet. *Am J Physiol Renal Physiol* 316:F558-f571. 10.1152/ajprenal.00474.2018

452 Opazo-Ríos L, Mas S, Marín-Royo G, Mezzano S, Gómez-Guerrero C, Moreno JA, and Egido J. 2020. Lipotoxicity
453 and Diabetic Nephropathy: Novel Mechanistic Insights and Therapeutic Opportunities. *Int J Mol Sci* 21.
454 10.3390/ijms21072632
455 Pepino MY, Kuda O, Samovski D, and Abumrad NA. 2014. Structure-function of CD36 and importance of fatty acid
456 signal transduction in fat metabolism. *Annu Rev Nutr* 34:281-303. 10.1146/annurev-nutr-071812-161220
457 Rego SM, and Snyder MP. 2019. High Throughput Sequencing and Assessing Disease Risk. *Cold Spring Harb*
458 *Perspect Med* 9. 10.1101/cshperspect.a026849
459 Sabiha B, Bhatti A, Roomi S, John P, and Ali J. 2021. In silico analysis of non-synonymous missense SNPs (nsSNPs)
460 in CPE, GNAS genes and experimental validation in type II diabetes mellitus through Next Generation
461 Sequencing. *Genomics* 113:2426-2440. 10.1016/j.ygeno.2021.05.022
462 Samsu N. 2021. Diabetic Nephropathy: Challenges in Pathogenesis, Diagnosis, and Treatment. *Biomed Res Int*
463 2021:1497449. 10.1155/2021/1497449
464 Tanase DM, Gosav EM, Neculae E, Costea CF, Ciocoiu M, Hurjui LL, Tarniceriu CC, Maranduca MA, Lacatusu CM,
465 Floria M, and Serban IL. 2020. Role of Gut Microbiota on Onset and Progression of Microvascular
466 Complications of Type 2 Diabetes (T2DM). *Nutrients* 12. 10.3390/nu12123719
467 Wang S, Dougherty EJ, and Danner RL. 2016. PPARγ signaling and emerging opportunities for improved
468 therapeutics. *Pharmacol Res* 111:76-85. 10.1016/j.phrs.2016.02.028
469 Yang C, Wang H, Zhao X, Matsushita K, Coresh J, Zhang L, and Zhao MH. 2020. CKD in China: Evolving Spectrum
470 and Public Health Implications. *Am J Kidney Dis* 76:258-264. 10.1053/j.ajkd.2019.05.032
471 Zhou F, and You G. 2007. Molecular insights into the structure-function relationship of organic anion transporters
472 OATs. *Pharm Res* 24:28-36. 10.1007/s11095-006-9144-9
473 Zhou H, Yu B, Sun J, Liu Z, Chen H, Ge L, and Chen D. 2021. Short-chain fatty acids can improve lipid and glucose
474 metabolism independently of the pig gut microbiota. *J Anim Sci Biotechnol* 12:61. 10.1186/s40104-021-
475 00581-3
476

Figure 1

The workflow of this study.

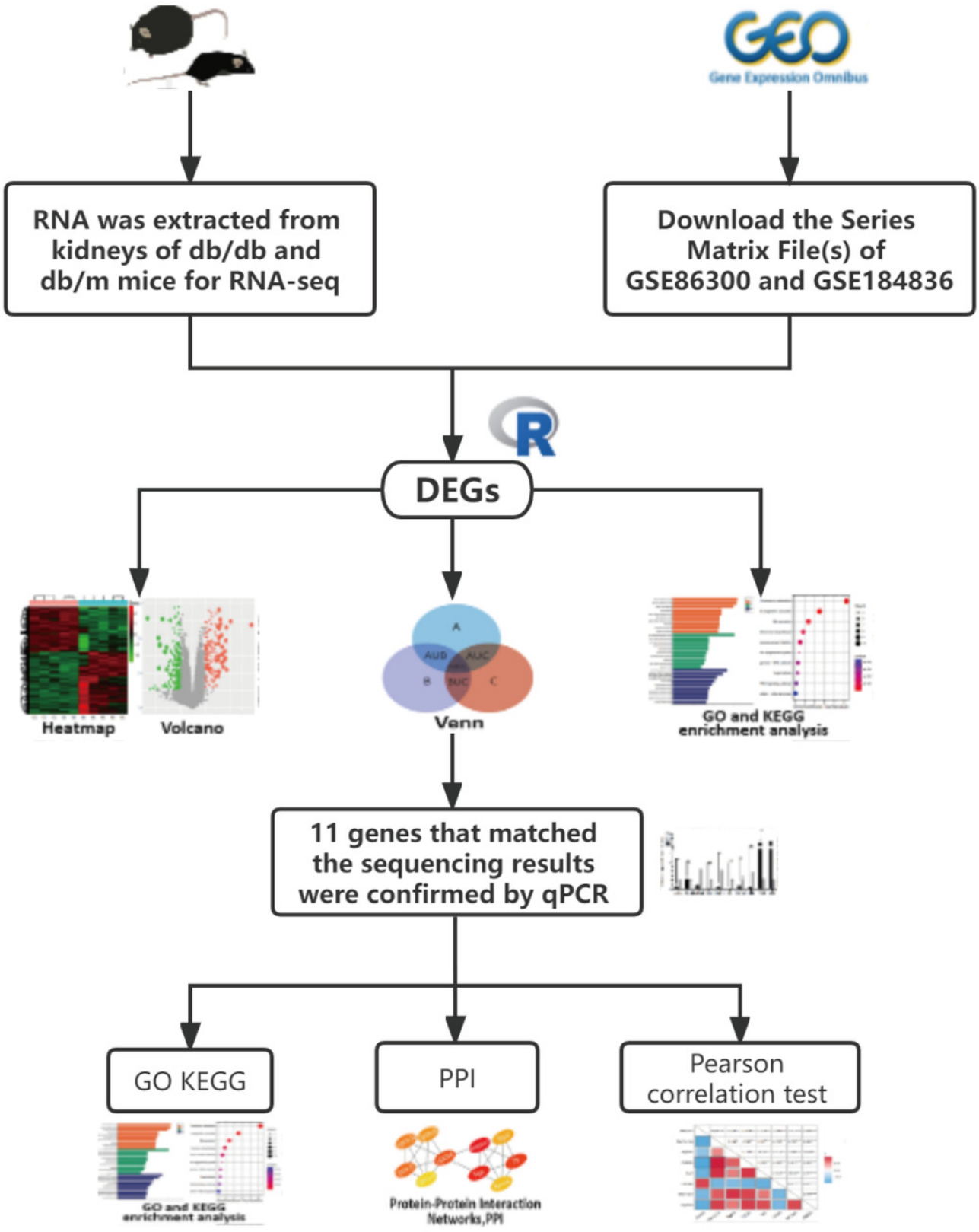
Gene Expression Omnibus
RNA was extracted from kidneys of db/db and db/m mice for RNA-seq
Download the Series Matrix File(s) of GSE86300 and GSE184836
DEGs
Heatmap
Volcano
Venn
GO and KEGG enrichment analysis
11 genes that matched the sequencing results were confirmed by qPCR
GO KEGG
PPI
Pearson correlation test
GO and KEGG enrichment analysis
Protein-Protein Interaction Networks,PPI

Figure 2

Biochemical parameters of blood and urine in DKD mouse model of different week ages

(a)Comparison of serum creatinine in mice; (b) Comparison of blood urea nitrogen in mice; (c)Comparison of total cholesterol in mice; (d) Comparison of triglycerides in mice; (e) Comparison of low-density lipoprotein in mice; (f)Comparison of high-density lipoprotein in mice; (g)Weight comparison of mice; (h) Comparison of blood glucose in mice; (i) Comparison of urinary albumin excretion rate (UAER) in mice. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

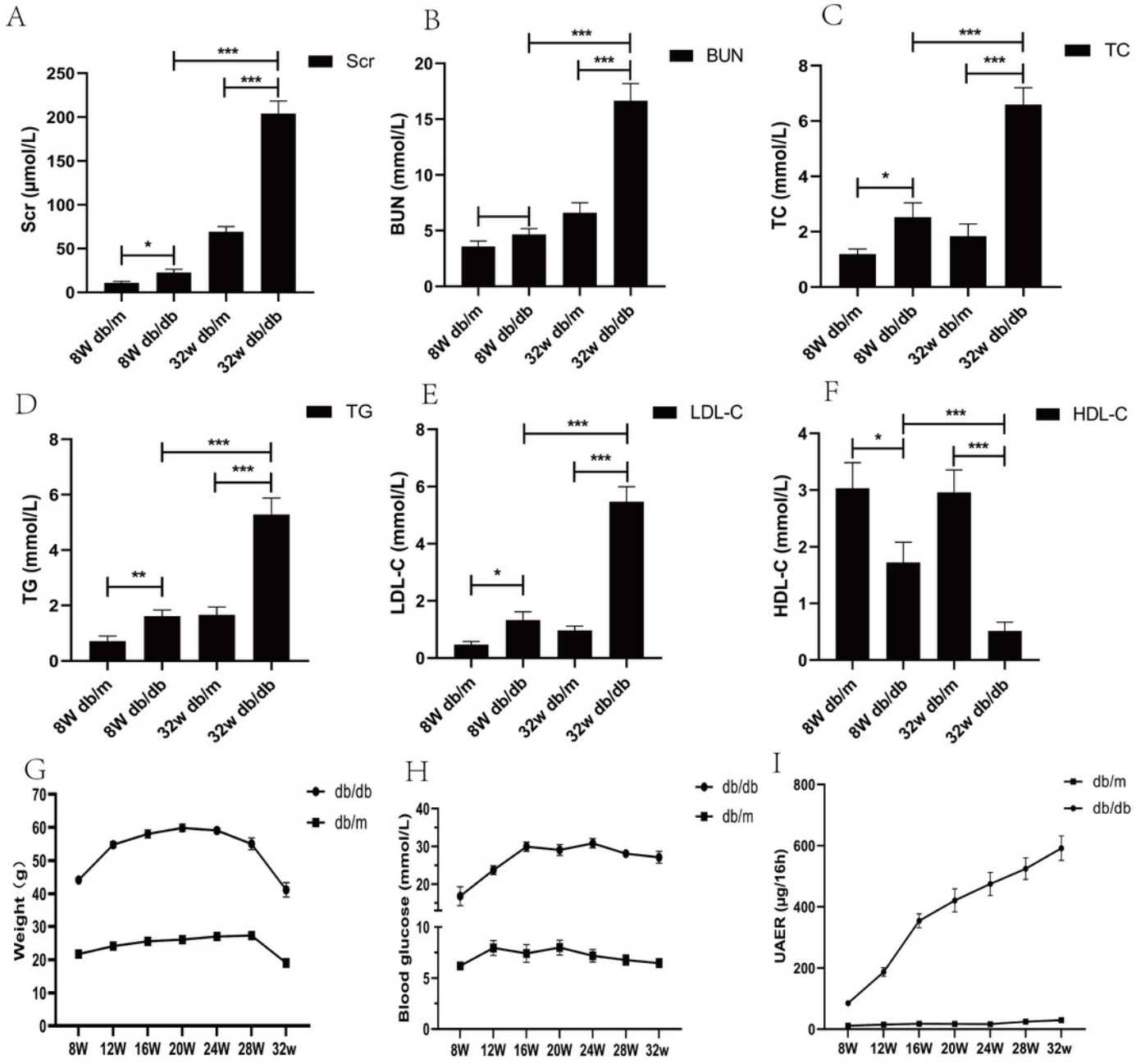
A
Scr
Scr (μmol/L)
B
BUN
BUN (mmol/L)
C
TC
TC (mmol/L)
D
TG
TG (mmol/L)
E
LDL-C
LDL-C (mmol/L)
F
HDL-C
HDL-C (mmol/L)
8W db/m
8W db/db
32w db/m
32w db/db
G
Weight (g)
H
Blood glucose (mmol/L)
I
UAER (μg/16h)
db/db
db/m
8W 12W 16W 20W 24W 28W 32w

Figure 3

Differential expression analysis of our RNA-seq and two GEO datasets (GSE86300 and GSE184836).

(a) Heatmap of DEGs in our RNA-seq. (b) Volcano map of our RNA-seq. A total of 59 upregulated and 66 downregulated DEGs were identified between the DKD and the Control group. (c) Heatmap of DEGs in GSE86300. (d) Volcano map of DEGs in GSE86300. A total of 183 up-regulated and 153 down-regulated DEGs were identified between the DKD and the Control group. (e) Heatmap of DEGs in GSE184836. (f) Volcano map of DEGs in GSE184836. A total of 76 up-regulated and 117 down-regulated DEGs were identified between the DKD and the control group. (g) Venn diagram of three DEGs groups. A total of 13 co-expression genes were obtained. Volcano map exhibit significantly differentially expressed genes, in volcano map, red bubbles mean up-regulated genes, blue bubbles mean down-regulated genes, and gray bubbles mean non-significant genes. The dots in the area above the horizontal dotted line have a P-value <0.05. The dots outside the two vertical dotted lines have a |log2FC| >1. Based on gene expression matrix, clustering analysis was shown in heatmap, in heatmap, red mean up-regulated genes, blue mean down-regulated genes. (|log2 FC|□1 and p-value □0.05)

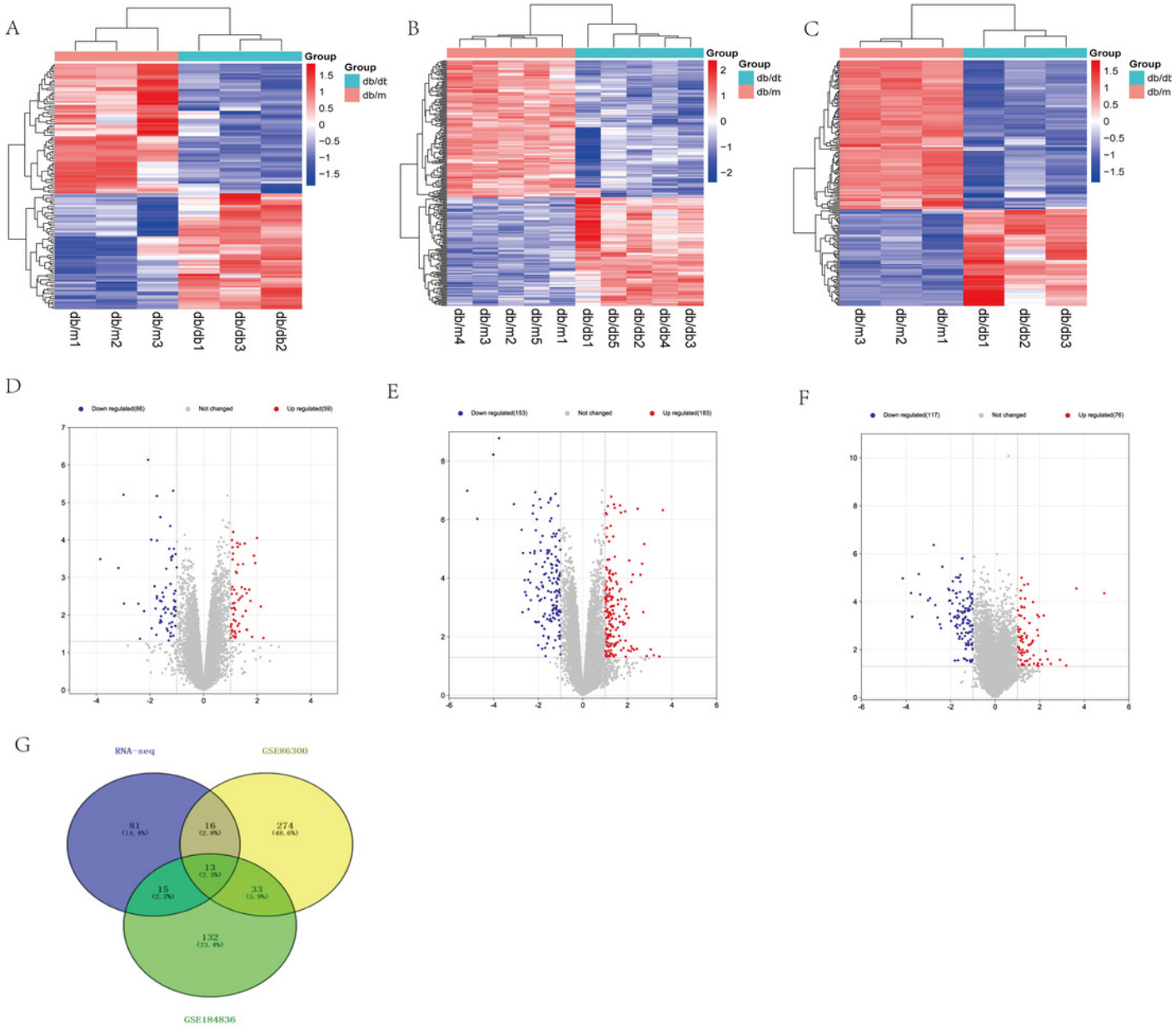
A
Group
db/dt
db/m
db/m1
db/m2
db/m3
db/db1
db/db3
db/db2
B
db/m4
db/m3
db/m2
db/m5
db/m1
db/db1
db/db5
db/db2
db/db4
db/db3
C
db/m3
db/m2
db/m1
db/db1
db/db2
db/db3
D
Down regulated(66)
Not changed
Up regulated(59)
E
Down regulated(153)
Not changed
Up regulated(183)
F
Down regulated(117)
Not changed
Up regulated(76)
G
RNA−seq
GSE86300
GSE184836
81
(14.4%)
16
(2.8%)
274
(48.6%)
13
(2.3%)
15
(2.7%)
33
(5.9%)
132
(23.4%)

Figure 4

Bar graph of GO Annotation and dot plot of KEGG pathway enrichment analysis of DEGs.

(a, b) Our RNA-seq (Top 10). (c, d) GSE86300 (Top 10). (e, f) GSE184836 (Top 10). Bar graph shows that DEGs of the three groups are enriched in several biological processes (BP), cell components (CC), molecular functions (MF). In the bar graph, we sorted the top 10 of BP, CC and MF by p-value and visualize them. In the dot plot, the color represents the p-value, and the size of the spots represents the gene number.

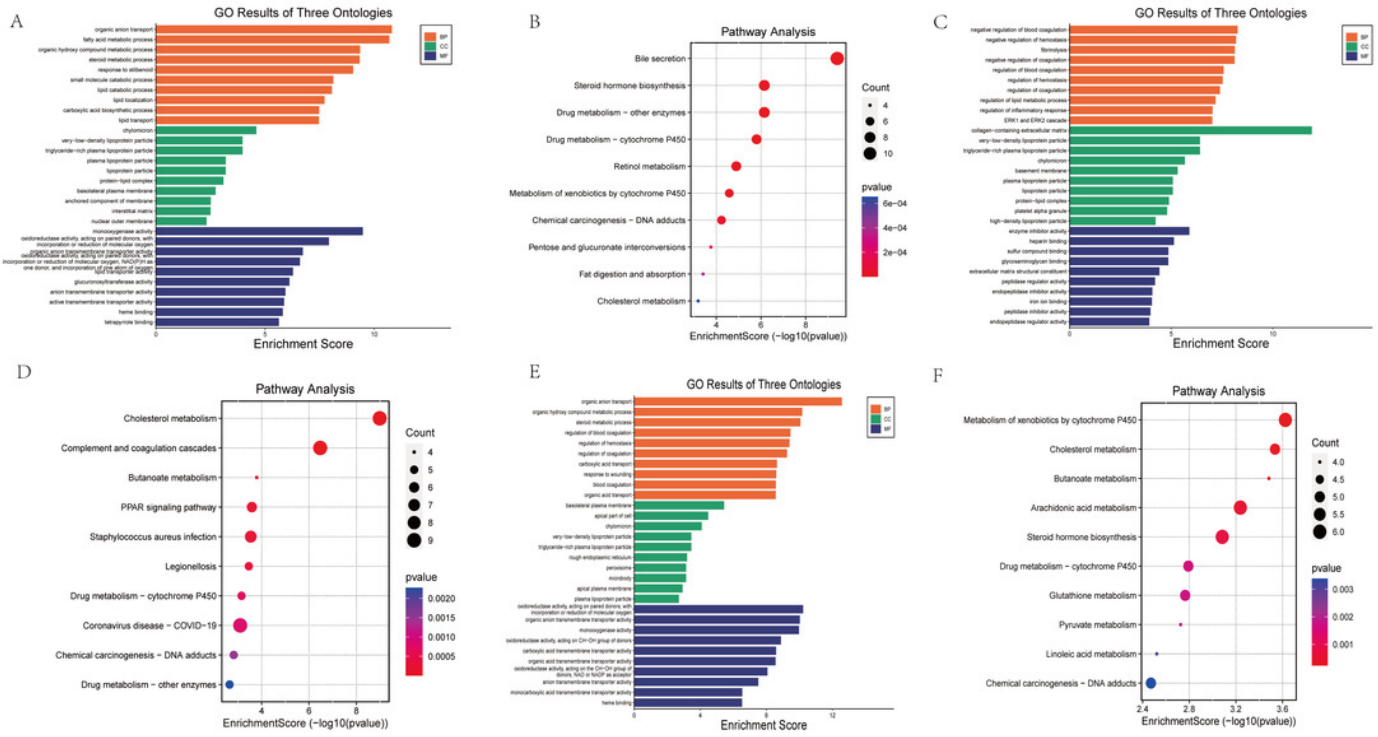

Figure 5

The relative mRNA expression of 13 overlapping genes in DKD and control group determined by qRT-PCR.

*p < 0.05, **p < 0.01, ***p < 0.001.

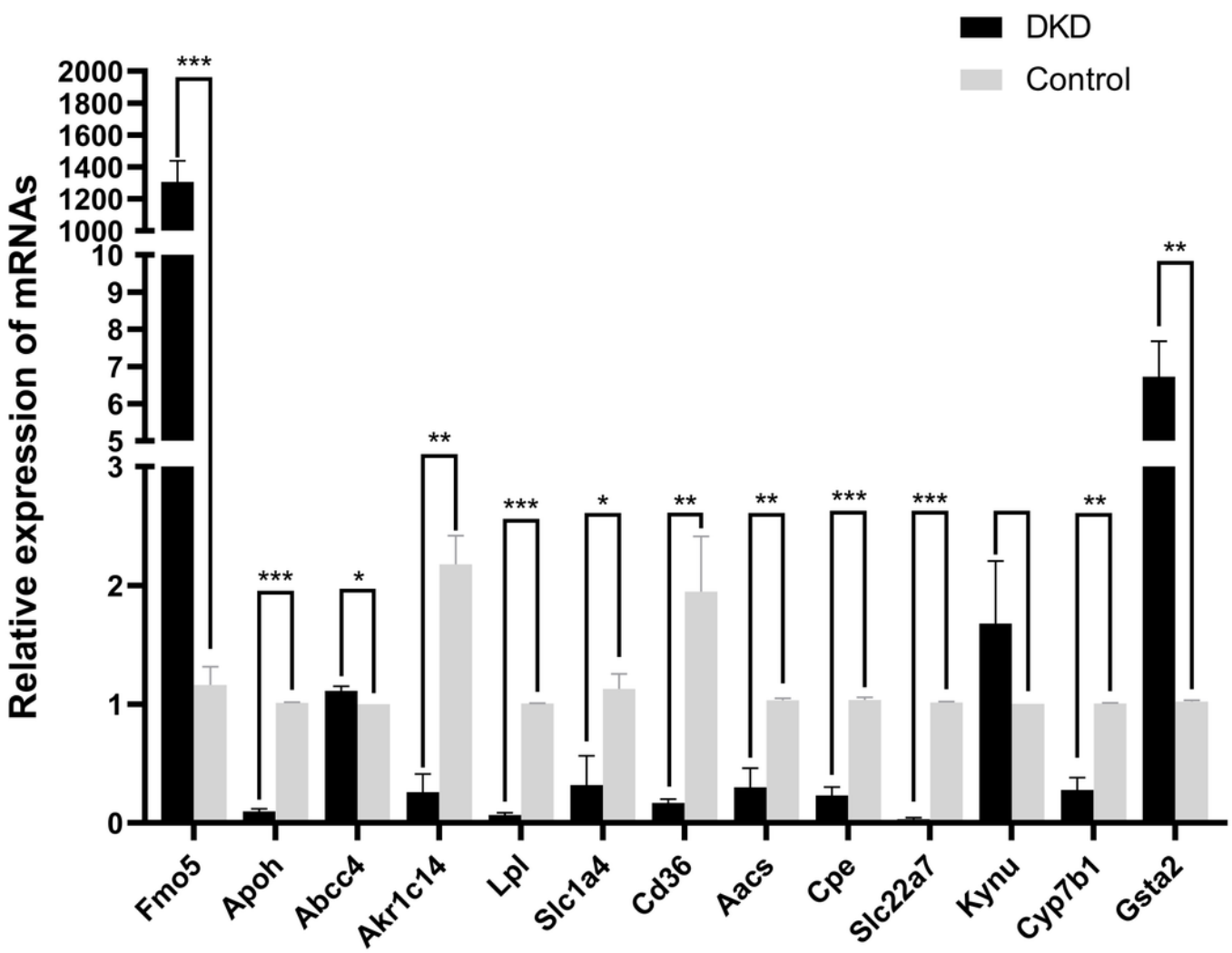

Figure 6

Functional analysis of key genes and their interactions.

(a) Dot plot of KEGG pathway enrichment analysis of verified 11 genes, in the dot plot, the color represents the p-value, and the size of the spots represents the gene number. (b) Bar plot of GO enrichment analysis of verified 11 genes. (c) PPI network of verified 11 DEGs, which also showed that other genes were involved. In the PPI analysis, green represents the monocarboxylic acid metabolic process, red represents the PPAR signaling pathway and blue represents the lipid metabolic process, white has no meaning. (d) The protein networks of the three key genes and several related genes. (e) Pearson correlation test between 11 DKD related genes, blue represents negative correlation, while red represents positive correlation.

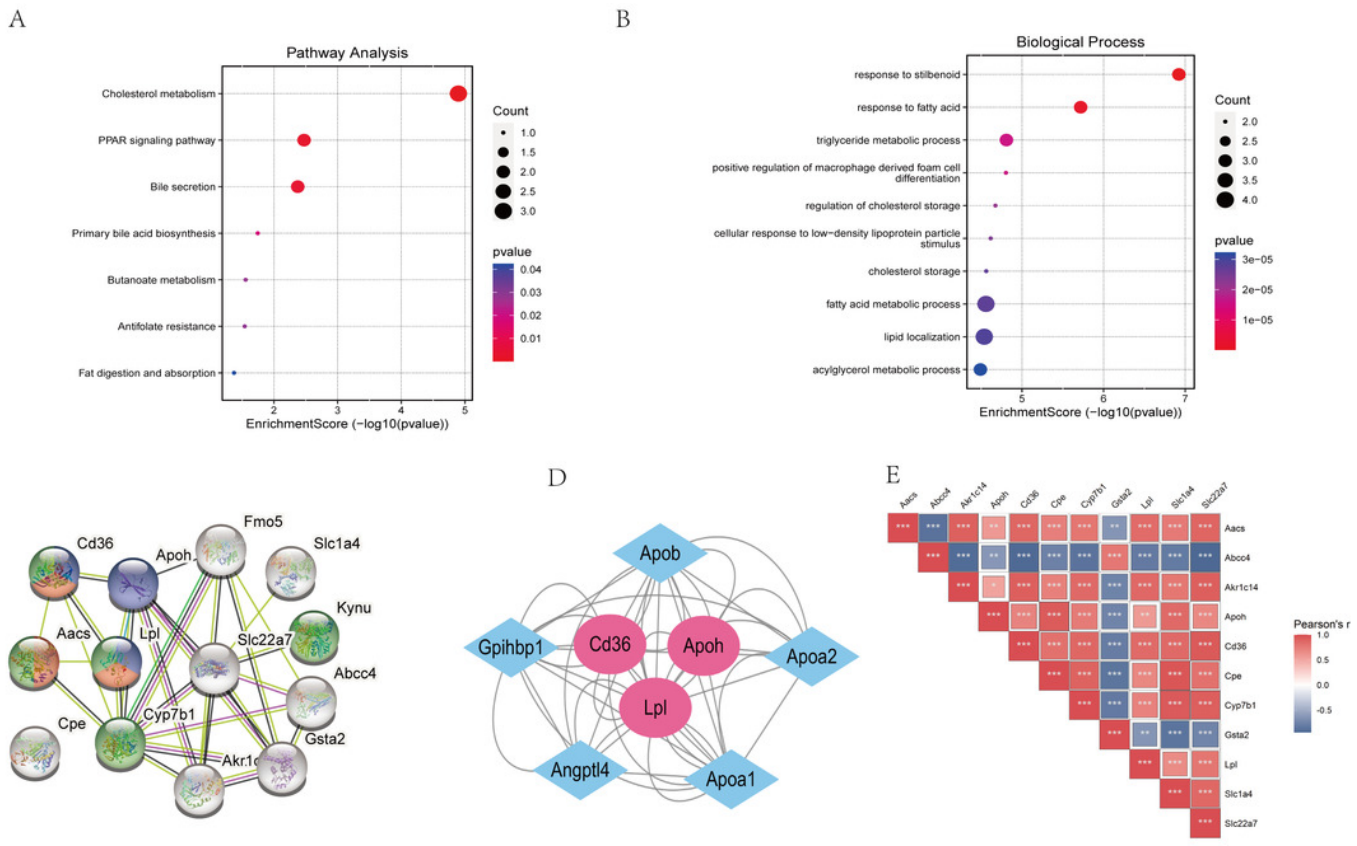

Table 1(on next page)

Expression of 13 overlapping genes in three datasets

1

group	RNA-seq		GSE86300		GSE184836		
Gene	logFC	p.Value	logFC	p.Value	logFC	p.Value	up/down
Aacs	-2.0679	0.0000	-1.3349	0.0016	-1.5640	0.0000	Down
Abcc4	1.1880	0.0145	1.4252	0.0000	1.3084	0.0002	Up
Akr1c14	-2.4325	0.0050	-1.8565	0.0001	-3.4254	0.0000	Down
Apoh	-1.4246	0.0087	-1.5422	0.0112	-1.1246	0.0002	Down
Cd36	-1.1073	0.0002	-2.1099	0.0000	-2.3700	0.0000	Down
Cpe	-1.7440	0.0000	-1.4616	0.0025	-2.0109	0.0001	Down
Cyp7b1	-2.9671	0.0050	-2.5894	0.0003	-3.0130	0.0001	Down
Fmo5	-1.5849	0.0108	-2.3630	0.0000	-1.4042	0.0014	Down
Gsta2	1.1099	0.0364	1.0850	0.0106	3.6420	0.0000	Up
Kynu	1.0834	0.0049	1.8903	0.0005	1.4218	0.0027	Up
Lpl	-1.0558	0.0023	-1.0108	0.0014	-1.6906	0.0022	Down
Slc1a4	-1.2027	0.0003	-1.2780	0.0010	-1.4501	0.0005	Down
Slc22a7	-3.8535	0.0003	-5.1868	0.0000	-4.1435	0.0000	Down

2

Table 2(on next page)

The BP analysis of 11 verified genes

1

ID	Description	*P*.value	Gene
GO:0035634	response to stilbenoid	1.20E-07	Cd36/Slc22a7/Gsta2
GO:0070542	response to fatty acid	1.91E-06	Aacs/Cd36/Lpl
GO:0006641	triglyceride metabolic process	1.55E-05	Cd36/Lpl/Apoh
GO:0010744	positive regulation of macrophage derived foam cell differentiation	1.57E-05	Cd36/Lpl
GO:0010885	regulation of cholesterol storage	2.12E-05	Cd36/Lpl
GO:0071404	cellular response to low-density lipoprotein particle stimulus	2.42E-05	Cd36/Lpl
GO:0010878	cholesterol storage	2.74E-05	Cd36/Lpl
GO:0006631	fatty acid metabolic process	2.76E-05	Aacs/Cd36/Lpl/Akr1c14
GO:0010876	lipid localization	2.90E-05	Cd36/Lpl/Apoh/Abcc4
GO:0006639	acylglycerol metabolic process	3.23E-05	Cd36/Lpl/Apoh

2